



ELSEVIER

Contents lists available at ScienceDirect

## Journal of Theoretical Biology

journal homepage: [www.elsevier.com/locate/yjtbi](http://www.elsevier.com/locate/yjtbi)

## Hidden patterns of reciprocity



Syi\*

Department of Sociology and Communication, Budapest University of Technology and Economics, Hungary

## HIGHLIGHTS

- Reciprocity can help the evolution of cooperation.
- In the context of direct reciprocity there exist four second-order action rules which are able to promote cooperation.
- In the context of indirect reciprocity there exist four second-order assessment rules which are able to promote cooperation.
- The four action rules and the four assessment rules can be paired, and they show very similar patterns.
- These common patterns are based on the relationship to the punishment.

## ARTICLE INFO

## Article history:

Received 24 August 2013

Received in revised form

14 November 2013

Accepted 9 December 2013

Available online 22 December 2013

## Keywords:

Cooperation

Evolutionary games

Direct reciprocity

Indirect reciprocity

Punishment

## ABSTRACT

Reciprocity can help the evolution of cooperation. To model both types of reciprocity, we need the concept of strategy. In the case of direct reciprocity there are four second-order action rules (Simple Tit-for-tat, Contribute Tit-for-tat, Pavlov, and Grim Trigger), which are able to promote cooperation. In the case of indirect reciprocity the key component of cooperation is the assessment rule. There are, again, four elementary second-order assessment rules (Image Scoring, Simple Standing, Stern Judging, and Shunning). The eight concepts can be formalized in an ontologically thin way we need only an action predicate and a value function, two agent concepts, and the constant of goodness. The formalism helps us to discover that the action and assessment rules can be paired, and that they show the same patterns. The logic of these patterns can be interpreted with the concept of punishment that has an inherent paradoxical nature.

© 2013 Elsevier Ltd. All rights reserved.

Human societies are based on cooperation. The big question is what kind of mechanism or strategy can provide and maintain the mutual cooperation among individuals when there is a strong temptation on mutual defection. These situations can be modeled by Prisoners's Dilemma (PD) or Iterated PD (IPD) game (Axelrod, 1984). Reciprocity is a very important concept of social sciences, especially in sociology and anthropology (see the works of Mauss, 1990; Malinowski, 1950; Sahlin, 1972), but it is used within the evolutionary biology as well (Trivers, 1971; Alexander, 1987). In the last two decades reciprocity was explored by game theory, and many theories and simulations were born (Sigmund, 2010). Trivers introduced the concept of direct reciprocity (DR) as a mechanism promoting and maintaining cooperation between two players who know each other. The latter moment, the expectation of familiarity has a natural condition: direct reciprocity provides cooperation only among small numbers of somehow related, at

least familiar individuals. The key component of the explanation is the action rule (ACR), and the key feature of the successful ACRs is the repetition (Sigmund, 2010).

Alexander has proposed that large-scale cooperation among humans can be explained with the help of the concept of indirect reciprocity (IR). After some early attempts (Boyd and Richerson, 1989) in 1998 Sigmund and Nowak developed a game theoretic formal model of indirect reciprocity. In order to explain IR we need to add a new component to the model of IR, an assessment rule (ASR). With the help of an ASR group members can change the reputation of individuals. In the context of indirect reciprocity players do not know each other, they do not know how the others behaved in the previous round and when individuals must decide about cooperation they can base their decision only on reputation information. There exists a very strong condition of the theories of reciprocity: actors have only two possibilities, they can only cooperate or defect. We can differentiate between two types of reciprocity: the one is positive reciprocity when the agent reciprocates something *Good* (Cooperation), the other is negative reciprocity when the agent repays something *Bad* (Defection).

\* Tel: +36 303116085.

E-mail address: [i@syi.hu](mailto:i@syi.hu)

### 1. Direct reciprocity

The language I use is an ordinary first-order predicate logic with two complementing deontic logic operators. Formalizing the concept of ACR first we need is an action predicate:  $do(agent, a, t_i)$  (Von Wright, 1963; Kanger and Kanger, 1966; Belnap et al., 2001). Here the category of action should have at least three parameters. First, there must be an *agent* who does or forbears something. In the context of reciprocity an *agent* can be *Ego* or *Alter*. The second parameter of the predicate refers to the content of the action, and the third parameter indicates the round of the repeated game environment. An action rule determines how the *agent* should decide for each round in an iterated game. In a second step we introduce a valuation/assessment function ( $VALUE(x, t)$ ), supposing a very simple completely Black & White world where things ( $x$ ) at a given moment ( $t$ ) can be assessed only as *Good* or *Bad* (where *Good* and *Bad* are undefined constants of our language)

$$G(x, t_i) \leftrightarrow VALUE(x, t_i) = Good$$

$$B(x, t_i) \leftrightarrow VALUE(x, t_i) = Bad$$

The two statements above can be expressed in another form. In a dichotomous, Black & White world it is true that

$$\neg(VALUE(x, t_i) = Good) \leftrightarrow VALUE(x, t_i) = Bad$$

or shortly

$$B(x, t_i) \leftrightarrow \neg G(x, t_i)$$

In the theories of reciprocity there are only two possible assessments of action: it can be *Good* or *Bad*. We can formalize it in the following way:

$$DO(agent, a, t_i) \rightarrow (VALUE(a, t_i) = Good \leftrightarrow \neg(VALUE(a, t_i) = Bad))$$

or shortly

$$DO(agent, a, t_i) \rightarrow (G(a, t_i) \leftrightarrow \neg B(a, t_i))$$

Based on the equations above we can define the category of cooperation ( $C(agent, t_i)$ ), and the category of defection ( $D(agent, t_i)$ )

$$C(agent, t_i) \leftrightarrow DO(agent, a, t_i) \wedge VALUE(a, t_i) = Good$$

$$D(agent, t_i) \leftrightarrow DO(agent, a, t_i) \wedge VALUE(a, t_i) = Bad$$

The connection between the two types of action is obvious

$$D(agent, t_i) \leftrightarrow \neg C(agent, t_i)$$

The next question is what kind of strategies (action rules) can be found, and how they can be formalized. There are conditional and unconditional ACRs. An example of an unconditional strategy is *AllD*, the so-called always defector (willing defector) strategy. Similarly, for *AllC* we can use an always cooperator (willing cooperator) strategy. The unconditional ACRs are too simple, they could not win a game series. Theoretically the condition of a conditional ACR can be any state of affairs, but in this context the conditional ACRs' condition is always a former action. It has two consequences: first, the extension of the concept of condition is narrower here as it is usual; second, the category of conditional action rule is a higher-order concept. If we would like to grasp the prescriptive character of our action rules we have to apply the *obligatory* operator from the field of deontic logic. When it is obligatory to do something for an agent this fact can be expressed in the following way:

$$O_{DO}(agent, a, t_i)$$

As a first step a simple first round sub-rule can be stated for all ACRs. In the first round cooperation is obligatory for both players.

This is an exception rule, because it refers only the first step

$$ACR(Ego, t_1) \leftrightarrow OC(Ego, t_1)$$

It is true for all action rules, so it is unnecessary to take into consideration any further. The four action rules can be interpreted without this sub-rule. Let us see how.

(i) Maybe the most famous experiment was Axelrod's round-robin tournament (Axelrod, 1984), where different strategies competed in an Iterated Prisoner's Dilemma game (IPD). The winning rule was the well-known *Tit-for-tat* (*tft*) strategy proposed by Rapoport. This rule cooperates in the first step, and in all other rounds repeats its partner's previous action. This is the prototypical ACR of direct reciprocity. *Tit-for-tat* strategy is very old, it can be found everywhere in our history. We have lots of proverbs with the same (or similar) meaning. "An eye for an eye, a tooth for a tooth." or "He who greets with a stick, will be answered with a club." is for the negative reciprocity, "One good turn deserves another." for positive reciprocity, and "Bread borrowed should be returned." for both. Axelrod (1984) evaluated *TFT* as a nice, retaliating, forgiving, and non-envious strategy. The next simple formula shows how we can describe the main rule of reciprocity: 'replicate your partner's moves'

$$TFT(Ego, t_i) \leftrightarrow ((D(Alter, t_{i-1}) \rightarrow OD(Ego, t_i)) \wedge (C(Alter, t_{i-1}) \rightarrow OC(Ego, t_i)))$$

This is very simple, partly because it is first-order rule. Ego's action exclusively depends on Alter's action in the previous round. But the all other action rules are higher-order, which has an important consequence: these action rules depend on the actions of both players in the previous round. In order to compare our formulas with each other we have to convert the *TFT*'s formula into new—redundant—form

$$TFT(Ego, t_i) \leftrightarrow ((C(Ego, t_{i-1}) \rightarrow (D(Alter, t_{i-1}) \rightarrow OD(Ego, t_i)) \wedge (C(Alter, t_{i-1}) \rightarrow OC(Ego, t_i))) \wedge (D(Ego, t_{i-1}) \rightarrow (D(Alter, t_{i-1}) \rightarrow OD(Ego, t_i)) \wedge (C(Alter, t_{i-1}) \rightarrow OC(Ego, t_i))))$$

(ii) Sugden (1986) has proposed a modified version of *Tit-for-tat*. He referred to it as  $T_1$ . Later it has been called *Standing* strategy in the field of indirect reciprocity, but Boyd labeled Sugden's strategy to *Contribute Tit-for-tat* ( $CTFT$ ) action rule (Boyd, 1989; Panchanathan and Boyd, 2003). Others called this rule as *Firm-But-Fair* (FBF) strategy (Frean, 1994; Hauert and Schuster, 1998). Although the 'firm, but fair' is a familiar, everyday life expression, I prefer the usage of the contrition-related term (Boerlijst et al., 1997). This rule can be characterized by a typical contrite attitude: if the player defected in the previous round, then the strategy prescribes unconditional cooperation. The fault must be corrected. The formula of this strategy is

$$CTFT(Ego, t_i) \leftrightarrow ((C(Ego, t_{i-1}) \rightarrow (D(Alter, t_{i-1}) \rightarrow OD(Ego, t_i)) \wedge (C(Alter, t_{i-1}) \rightarrow OC(Ego, t_i))) \wedge (D(Ego, t_{i-1}) \rightarrow OC(Ego, t_i)))$$

This is the most cooperative strategy, it subscribes cooperation in three cases (from the four possibilities). Probably all stories about contrition can be related to  $CTFT$ . One of the parables of Jesus describes the same rule: "He that is without sin among you, let him first cast a stone at her." (John 8.1-11).

(iii) Independent from the theories of reciprocity Kraines and Kraines (1989, 1993) introduced and analyzed the Pavlov strategy. Rapoport called it *Simpleton* rule (Ridley, 1997), some authors use the *Perfect TFT* name for it (Imhofa et al., 2007), and with the emergence of game theoretic modeling a new term, *Win-Stay-Lose-Shift* (WSLS) appeared on the scene. It is a kind of learning strategy and maybe the most successful rule that outperforms *Simple TFT* in a noisy environment where social error exists

(Nowak and Sigmund, 1993; Imhofa et al., 2007). The formula of this ACR is

$$\begin{aligned} \text{PAVLOV}(Ego, t_i) \leftrightarrow & ((C(Ego, t_{i-1}) \rightarrow (D(Alter, t_{i-1}) \\ & \rightarrow \mathbf{O}_D(Ego, t_i)) \wedge (C(Alter, t_{i-1}) \rightarrow \mathbf{O}_C(Ego, t_i))) \wedge (D(Ego, t_{i-1}) \\ & \rightarrow (C(Alter, t_{i-1}) \rightarrow \mathbf{O}_D(Ego, t_i)) \wedge (C(Alter, t_{i-1}) \rightarrow \mathbf{O}_D(Ego, t_i)))) \end{aligned}$$

The formula shows that this is the most complex strategy. If *Ego* cooperated in the previous round, then the first subrule prescribes to replicate *Alter* previous move, but if *Ego* defected in the previous round, then the second subrule prescribes an opposite replication. In that case *Alter* has to do the complement of what *Alter* did in the previous round. On the other hand this complexity can—at least partly—explain why this strategy performs so successfully.

(iv) The fourth ACR is the *Grim Trigger* ( $\text{GT}$ ) action rule (Friedman, 1971). Sometimes it is called Friedman or permanent retaliation strategy. This is an unforgiving, resentful strategy. Why? It cooperates as long as its partner cooperates, but after the first defection of the opponent this strategy switches to an endless defection. Just one mistake, one defection is enough: there is no forgiveness. Its formula

$$\begin{aligned} \text{GT}(Ego, t_i) \leftrightarrow & ((C(Ego, t_{i-1}) \rightarrow (D(Alter, t_{i-1}) \\ & \rightarrow \mathbf{O}_D(Ego, t_i)) \wedge (C(Alter, t_{i-1}) \rightarrow \mathbf{O}_C(Ego, t_i))) \wedge (D(Ego, t_{i-1}) \\ & \rightarrow \mathbf{O}_D(Ego, t_i))) \end{aligned}$$

The rule has no well-known everyday life representation (probably because of its high defection ratio), but some scholars supposed that  $\text{GT}$  strategy can be successful: “The so-called folk theorem on repeated games implies that if the probability for future rounds is sufficiently high, cooperation can be sustained by so-called trigger strategies, which switch to relentless defection as soon as the co-player defects once.” (Nowak and Sigmund, 2005)

(v) These four direct reciprocity action rules were on the focus of the game theoretic analysis. If we represent them together in a table, we can discover some interesting common features. The two *agents* have two possibilities (*C* or *D*), so in every round there are four possible outputs of an ACR: mutual cooperation (CC), mutual defection (DD), and two opposite actions (CD and DC). The summary of four strategies can be seen in Table 1.

The common feature of the four ACRs is that if *Ego* cooperated in the previous round, then all rules require the exact replication of *Alter*’s previous step. The difference between the four strategies can be found when *Ego* defected earlier. Theoretically there are only four possibilities for *Ego* to defect. We have four ACRs, and they cover these four theoretical alternatives. It means that the four ACRs are jointly exhausted and mutually incompatible. There is nothing more, no other rule.

## 2. Indirect reciprocity

Indirect reciprocity (IR) is the mechanism of large-scale cooperation. Sigmund and Nowak (1998) and Nowak and Sigmund (1998) proposed the first formal model for IR. In this model the

**Table 1**  
Summary of the four ACRs.

ACRs↓	$\text{do}(Ego, a, t_{i-1}) \rightarrow$	C		D	
			D	C	D
Simple Tit-for-tat	$\mathbf{O}_{\text{do}}(Ego, a, t_i) \Leftrightarrow$	C	D	C	D
Contrite Tit-for-tat		C	D	C	C
Pavlov		C	D	D	C
Grim Trigger		C	D	D	D

concept of strategy must be expanded. While in the case of DR the notion of strategy is synonymous with the concept of action rule (ACR), in the new context we must add the category of assessment rule (ASR) to the model of IR. Hereafter the strategy has two components (or modules): an action rule and an assessment rule. An ASR determines the reputation of an agent. Reputation is a special value concept that refers only agents (it has narrower range than the general  $\text{value}(x, t_i)$  concept). Reputation has two possible types:

$$\begin{aligned} G(\text{agent}, t_i) \leftrightarrow & \text{VALUE}(\text{agent}, t_i) = \text{Good} \\ B(\text{agent}, t_i) \leftrightarrow & \text{VALUE}(\text{agent}, t_i) = \text{Bad} \end{aligned}$$

Although an indirect reciprocity strategy has two components, an action rule and an assessment rule, here the former has no real importance, so it can be omitted (of course a more detailed analysis would require to take it into account). Hereafter, we deal only with ASRs. The key question is how the *agents*’ reputations can be determined and how they can be changed. An assessment rule is a valuation method that describes (or more exactly prescribes) how the reputation of *Ego* (in this context: *Donor*) is (must be) changed after *Ego* (*Donor*) cooperated or defected against *Alter* (in this context: *Recipient*). Describing the assessment activity we need another deontic operator (**A**). For the sake of illustration when the community gives a new (*Bad*) reputation to an *agent*, we can describe it in this way

$$\mathbf{A}\{\text{VALUE}(\text{agent}, t_i) = \text{Bad}\}$$

(i) In their first formal model of IR Sigmund and Nowak (1998) introduced *Image Scoring* method, which is a very simple ASR. In the first round both players have *Good* reputation

$$\mathbf{A}\{\text{VALUE}(\text{agent}, t_1) = \text{Good}\}$$

From the second round reputation can be changed. After observing *Ego*’s action his/her reputation will be *Good*, if (s)he cooperated, and *Bad*, if s(he) defected. We can represent it with the following formula (using both the  $\text{value}(x, t_i)$  function and  $\text{do}(\text{agent}, a, t_i)$  predicate, and eliminating again the round-related conditions):

$$\text{DO}(Ego, a, t_i) \rightarrow \text{VALUE}(Ego, t_i) = \text{VALUE}(a, t_i)$$

Based on this connection we can describe the *Image Scoring* assessment rule

$$\begin{aligned} \text{IS}(Ego, t_i) \leftrightarrow & ((G(Alter, t_i) \rightarrow \mathbf{A}\{\text{DO}(Ego, a, t_i) \rightarrow \text{VALUE}(Ego, t_i) \\ & = \text{VALUE}(a, t_i)\}) \wedge (B(Alter, t_i) \rightarrow \mathbf{A}\{\text{DO}(Ego, a, t_i) \rightarrow \text{VALUE}(Ego, t_i) \\ & = \text{VALUE}(a, t_i)\})) \end{aligned}$$

The equation can be interpreted that the new reputation of *Ego* after the  $t_i$  round is always equal to the value of *Ego*’s action in the  $t_i$  round. The formula above reveals some important features of the *Image Scoring* method: it clearly shows that we need only to observe *Ego*’s action, and that this method is indifferent to the existence (action, reputation) of the other participant (*Alter*). It is elementary. *Ego*’s action can be easily observed, it is unnecessary to deal with motivations, this method requires the less information, and, finally, it seems to be very intuitive—at first sight. We should always cooperate, true? Not always. This simplicity can sometimes be hypocrisy. It is perhaps no coincidence that Jesus’ Sermon on the Mount contains a lot of paragraphs against hypocrisy like this: “When you pray, you are not to be like the hypocrites; for they love to stand and pray in the synagogues and on the street corners so that they may be seen by men. Truly I say to you, they have their reward in full.”

(ii) The exciting question here is why we should cooperate with a *Bad* partner. If we punish *Bad* persons, it means, that we do not

**Table 2**  
Summary of the four ASRs.

ASRs↓	VALUE(Alter, t <sub>i</sub> ) → DO(Ego, a, t <sub>i</sub> ) →	G		B	
		C	D	C	D
Image Scoring	<b>A</b> {VALUE(Ego, t <sub>i</sub> ) = x} ⇔	G	B	G	B
Simple Standing		G	B	G	G
Stern Judging		G	B	B	G
Shunning		G	B	B	B

help them, so we defect, but this defection is justified and deserved. This principle gives the essence of the next ASR, the *Simple Standing* (ss) assessment rule. This method is similar to the *Image Scoring* in the case of a *Good* partner, but if *Alter* is *Bad* then this rule says that punishing the *Bad Recipient* is *Good*. These can be expressed with this formula

$$ss(Ego, t_i) \leftrightarrow ((G(Alter, t_i) \rightarrow \mathbf{A}\{DO(Ego, a, t_i) \rightarrow VALUE(Ego, t_i) = VALUE(a, t_i)\}) \wedge (B(Alter, t_i) \rightarrow \mathbf{A}\{G(Ego, t_i)\}))$$

This method already takes into account both the *Ego's* action and the partner's reputation. Handling the *Recipient's* reputation means that altruism manifested in cooperation should be discriminative (Hardin, 1982). This is the ideology (or practical reason) of deserved punishment. When we tell stories to our children, we frequently draw the conclusion: "He who is bad, gets his deserved punishment." Our historical experiences and the new game theoretic simulations as well verify the principle: if you do bad, you get punished (or punishing the wrongdoer must be rewarded). This is the logic of *Simple Standing* method (Ohtsuki and Iwasa, 2004, 2006, 2007).

(iii) But this is not the end. If we accept deserved punishment, we can take one more step: we can be inexorable. If we want to punish the *Bad* because we believe that we can prevent defections, we can ask ourselves: 'Is it good to cooperate with a bad person'. According to the *Simple Standing* method the answer is 'yes'. Another assessment rule says it is unacceptable. The name of the new method is *Stern Judging* (sj) (Pacheco et al., 2006; Brandt and Sigmund, 2004), but in an earlier paper Kandori (1992) described the same mechanism, and there are lots of references to *Kandori's* method as well. The formula of this rule is

$$sj(Ego, t_i) \leftrightarrow ((G(Alter, t_i) \rightarrow \mathbf{A}\{DO(Ego, a, t_i) \rightarrow VALUE(Ego, t_i) = VALUE(a, t_i)\}) \wedge (B(Alter, t_i) \rightarrow \mathbf{A}\{DO(Ego, a, t_i) \rightarrow VALUE(Ego, t_i) \neq VALUE(a, t_i)\}))$$

In the case of a *Good Recipient* this rule is the same as the other method, but for the *Bad* person it prescribes the opposite what the *Image Scoring* strategy specifies. This rule is really inexorable with the *Bad* people, but it leaves no rest for *Good* people, who will be punished if they cooperate with *Bad* individuals. In that case the argument is "Among the sinners the silent is guilty". But, we can recall an everyday life situation when after a kids' party seeing the messy room a mother wants to discipline her son and the boy defends himself "The others did it", the mother rebukes him: "Why did not you rebuke them?" This is the logic of *Stern Judging*: punishing who does not punish (the *Bad*).

(iv) This severity does not destroy the capacity of *Stern Judging* ASR promoting and maintaining cooperation. If we further reinforce our punishment policy, we start to threaten our results. The *Shunning* (sh) assessment method is the strongest penalty rule (sometimes it is called the strict-discriminator rule). This mechanism punishes the decision-holder player in three cases (Ohtsuki and Iwasa, 2007; Takahashi and Mashima, 2003; Panchanathan

and Boyd, 2004). This is its formula

$$sh(Ego, t_i) \leftrightarrow ((G(Alter, t_i) \rightarrow \mathbf{A}\{DO(Ego, a, t_i) \rightarrow VALUE(Ego, t_i) = VALUE(a, t_i)\}) \wedge (B(Alter, t_i) \rightarrow \mathbf{A}\{B(Ego, t_i)\}))$$

It is similar to the *Simple Standing* method, the only difference is that in the cases of the *Bad Recipient* the rule prescribes unconditional *Bad* assessment. The *Donor* can do anything, his/her reputation will always be *Bad*. This rule forces *agents* to shun *Bad* people: "Living with a devil, you have to turn a devil". It is not surprising that—according to the game theoretic simulations—this ASR does not really support cooperation. *Shunning* rule is the destructive logic of segregation.

(v) As we showed the four DR rules in one table, now we do it again. The summary of the four IR rules can be seen in Table 2.

The common feature of the four rules is that they evaluate in the same way how to behave against *Good Recipient*. The difference between them is based on how they assess *Ego's* actions against *Bad Recipient* (Rosas, 2010). But we have not yet addressed the question of why humans need different action and assessment rules.

Earlier I used the category of deserved (or justifiable) punishment but it is not clear why it is important. We can use the notion of constitutive (counts as) rule proposed by Searle (1969), to answer the question. What does it mean 'punishing' somebody? Doing something *Wrong* (*Bad action*, *Defection*) and saying 'this wrongdoing counts as a punishment'. We have to distinguish two different theses. (i) Punishment always refers to an earlier first-order level action, so it is a second-order concept. It causes the second-order Prisoner's Dilemma (Oliver, 1980; Boyd and Richerson, 1992). (ii) Punishing has always two evaluations. When we punish we do or forbear something (for example we impose a fine or forbear to donate), and we claim that it counts as punishment on the Searlian institutional level. And as is proper for a second-order notion at this point a paradox arises: it is never clear which interpretation is used when somebody evaluates something. Let us see why!

The second-order Prisoner's Dilemma arises when we incorporate punishment into the games. Punishment can promote cooperation but it is also costly. The costly punishment is collective goods, and the players feel strong temptation to defect, so a new question arises: who provides the costly punishment? From this point of view this type of situation becomes the Prisoner's Dilemma game—on the second-order level. If we suppose that on the first level cooperating means donating to somebody and defecting is not donating, on the second-order level, punishment is the cooperation that is equal to withdrawal of donation (not helping). If we represent it in a table we can see the paradox that I indicated earlier (Table 3).

The *paradox of punishment* is obvious. If we help we cooperate on first-order but defect on second-order. If we do not help we defect on first-order but cooperate on second-order. We always have to choose, and there are always more alternatives. That is why we need more evaluation systems. *Image Scoring* punishes the punishment, *Simple Standing* is the promoter of the first-order punishment, *Stern Judging* promotes the second-order punishment

**Table 3**  
The paradox of punishment.

	Second-order cooperation =Punishing =Not helping	Second-order defection =Not punishing =Helping
First-order cooperation =Helping	Contradiction	Helping
First-order defection =Not helping	Not helping	Contradiction

**Table 4**  
Decomposition of the eight third-order ASRs.

ACRs ↓	Good Ego	Bad Ego
Standing	Simple Standing	Image Scoring
Simple Standing	Simple Standing	Simple Standing
ACR <sub>1</sub>	Simple Standing	Stern Judging
Strict Standing	Simple Standing	Shunning
ACR <sub>2</sub>	Stern Judging	Image Scoring
ACR <sub>3</sub>	Stern Judging	Simple Standing
Stern Judging	Stern Judging	Stern Judging
Judging	Stern Judging	Shunning

(because it punishes abstaining from first-order punishment), and *Shunning* punishes not the action but the agent (because *Alter* can do anything, can cooperate or defect no matter, *Ego* punishes *Alter*, which means the reason of the punishment is not the action but the—bad—person).

(vi) The four ASRs in the table are secondary rules, because they have two parameters (*Ego*'s action and *Alter*'s reputation). We can refine our model if IR, if we add a new parameter, and take into account of the reputation of *Ego*. In that case we get third-order strategies. Theoretically there exist 4096 third-order IR strategies, but—according to the simulations—only eight from them are evolutionarily stable and can provide cooperative pattern (Ohtsuki and Iwasa, 2004, 2006). And what is very important: these eight third-order indirect assessment rules can be composed with the help of the four secondary ASRs analyzed above. If we separate the cases of *Bad Ego* and *Good Ego*, we get a very interesting pattern. Every third-order ASR can be decomposed into two second-order ASR (Table 4).

The first column shows the names of the strategies. Some of them are familiar, some are new. *Standing* or *Judging* is the third-order version of *Simple Standing* or *Stern Judging* strategies (Sugden, 1986). It can be seen that three of them have no given name (ACR<sub>i</sub>). We know that they exist, but we do not know their descriptions. If *Ego* is *Good*, then only two ARSs (the *Simple Standing* and the *Stern Judging* methods) can be applied, and in both cases the four possible second-order ASRs can be distributed. The conclusion is a little bit surprising: stepping up to the third-order level we need no further component, we can expand our model without any new ontological commitment. We can describe the whole phenomenon of indirect reciprocity based only on the four second-order ASRs. Based on the four “cooperation provider” second-order assessment rules we can compose 16 third-order assessment modules theoretically, but only half of them are among the leading eight. When *Ego* is *Good* only the *Simple Standing* and the *Stern Judging* rules can be applied. Why cannot the other two second-order assessment modules be used in that case? Before answering this we need some general remarks.

### 3. The common patterns of DR and IR

If we generalize further we can discover a hidden common pattern in the structure of the two types of reciprocity system. In the table of the direct reciprocity rules the four ACRs were represented by the type of *Ego*'s action (*C* or *D*), while in the other table the ASR methods were represented by the type of *Alter*'s reputation (*G* or *B*). If we change both value pairs to 1 and 0 values, we get the same patterns (Table 5).

We can identify four pairs among the eight DR and IR strategies. *Simple TFT* and *Image Scoring*, *Contrite TFT* and *Simple Standing*, *Pavlov* and *Stern Judging*, and finally *Grim Trigger* and *Shunning* are very close to each other (Rosas, 2010). The two reciprocity systems handle different information. DR strategies

**Table 5**  
The common patterns of the four strategy pairs.

ACRs ↓	C		D		← DO( <i>Ego</i> , <i>a</i> , <i>t<sub>i-1</sub></i> )
	C	D	C	D	← DO( <i>Alter</i> , <i>a</i> , <i>t<sub>i-1</sub></i> )
Simple TFT	1	0	1	0	Image Scoring
Contrite TFT	1	0	1	1	Simple Standing
Pavlov	1	0	0	1	Stern Judging
Grim Trigger	1	0	0	0	Shunning
DO( <i>Ego</i> , <i>a</i> , <i>t<sub>i</sub></i> ) →	C	D	C	D	↑ ASRs
VALUE( <i>Alter</i> , <i>t<sub>i</sub></i> ) →	G		B		

take into account the two players' actions in the previous round, IR strategies assess *Ego*'s action and *Alter*'s reputation in the given round. They differ in their outputs: DR rules prescribe actions, IR systems offer reputations for *Ego*. The two types of reciprocity rules evaluate the action of an agent (in DR *Alter*'s action, in IR *Ego*'s action) and the other information (*Ego*'s action or *Alter*'s reputation) provides the “context” of this evaluation. If we focus on the common value content of this information the common patterns can be understood more easily. The first—very strong—common rule refers to the state when the “things are good”, when the context of the evaluated agent is *Good* that is the other agent cooperates or the partner has good reputation. In that case the primary rule prescribes the repetition of the value content of the evaluated action in order to maintain the cooperation in the community. If the evaluated agent cooperated (defected) i.e. acted in *Good* (*Bad*) way then all rules prescribe something *Good* (*Bad*): in DR the cooperation (defection) for *Ego*, in IR *Good* (*Bad*) assessment of *Ego*. That is why the first two columns of Table 5 contain the same values for all rules. This primary rule is common to all eight reciprocal strategies. This is a *TFT*-like rule because the essence of the rule is the exact repetition of the value content of the action of the evaluated agent. But again, the primary rule is valid only in a narrower domain when the evaluation context is *Good*. In that case the evaluation state can be purely *Good* (*C* and *C* in DR, *G* and *C* in IR) or mixed (that is *C* and *D* in DR, *G* and *D* in IR). In the *Good* state the cooperation prescribed by the primary rule obviously maintains the cooperative state in the community. When the action of the evaluated agent is *Bad* (the evaluation state is mixed), the prescribed defection can be interpreted as a first-order punishment which can be enough to make the evaluated person redirect to the cooperative path.

The four strategy pairs differ from each other when the context of the evaluated action is *Bad* that is when *Ego* defected in the previous round (in DR) or *Alter*'s reputation is *Bad* (in IR). What changes when we consider that the context can be *Bad*? In that case the evaluation state can be mixed (*D* and *C* in DR, *B* and *C* in IR) or *Bad* (*D* and *D* in DR, *B* and *D* in IR). The latter case raises the most important question: how it is possible to return to the double *Good* state (*CC* in DR or *CG* in IR) from the state of the double *Bad* state (*DD* or *DB*). The answer can be found and the similarity between the examined strategies can be observed in their relationship to the punishment.

*Scoring* and *TFT* are both first-order rules. Their patterns and their roles in the reciprocity are the same: “We expect, however, that *Scoring* has a similar role in indirect reciprocity to that of *tit-for-tat* in direct reciprocity. Neither strategy is evolutionarily stable, but their ability to catalyze cooperation in adverse situations and their simplicity constitute their strength.” (Nowak, 2006). Both rules oppose the punishment, although in a hypocritical way. Their common principle is punish the punishment.

When Sugden (1986) proposed his *T<sub>1</sub>* (*Standing*) strategy, he did not talk about reputation, but he used a very similar concept to

reputation: 'standing'. Later the *Standing* (*Simple Standing*) rule became part of the IR theory as an assessment rule. When Boyd (1989) relabeled Sugden's strategy to *Contrite tit-for-tat*, he referred to it as an action rule. Is it an assessment or an action rule? Is not it a terminological confusion? No, rather, it shows the close connection between the two types of rule. When Frean proposed *Firm-But-Fair* label, he gave the following explanation: "... it is 'firm' in that it retaliates by defecting if it was a sucker in the previous round. It is also 'fair' in that it does not retaliate against a defector if it defected itself, and it cooperates with suckers rather than continuing to exploit them." (Frean, 1994). This argument is similar to what was said about the *Contrite TFT* rule, although the emphasis is shifting to the punishment. At first sight it can be strange to talk about punishment in the context of DR but, maybe, it can be acceptable if we examine what the concept of contrition means. After I do something wrong than I feel remorse and I would like to compensate for my "sin". How can I compensate in a DR situation? If I had no contrition my natural reaction would be defection against an agent who defected previously. If I had contrition and I would like to compensate for my defection in the previous round I would cooperate with *Alter* who defected in the previous round. This is a special kind of self-punishment when the agent who sanctions and the agent who is sanctioned are the actors itself. This punishment is directed to the contrite agent so it is internal, and in this sense we can say that this self-punishment is justified. In IR-context the punishment is external, and its justified character is obvious. When Weber (1978) talked about the types of legitimate norm (order) he also separated internal (subjective) and external solutions. From this point of view the objective of the external punishment is to make *Alter* feel contrition in the hope that (s)he will cooperate in the future. Their common slogan: do not punish the punishment.

The earlier description emphasized the learning capability of the *Pavlov* rule. Rosas (2010) evaluated *Pavlovian/Stern Judging* dual-rule as a flattery, an obsequious strategy. Focusing on the common punishment related component of this dual rule can be interpreted similarly to the *Contrite TFT/Simple Standing* strategy pair. The most important similarity between the two strategy pairs is that both prescribe to do something *Good* (*cooperation* in DR, *Good* reputation in IR). It means that *Pavlov* is the same contrite strategy as *CTFT*, and *Stern Judging* is the same justified punishment rule as *Simple Standing*. That is why these two strategy pairs can successfully promote cooperation in a community. These rules go from the double *Bad* state (*DD* or *DB*) to a *Good* one (*C* or *G*), so they can break the death-spiral of mutual defection. Probably this very important common property explains why the *Simple Standing* and the *Stern Judging* modules can only be applied in the case of *Good* donors during the decomposition of the leading eight third-order assessment rules and that can be the reason why we have only eight evolutionary cooperative modules from the theoretically possible sixteen. The *Shunning* and the *Image Scoring* rules cannot always provide stable cooperation and this is unacceptable in the case of *Good Donors* according to our primary rule.

But what is the difference between the *Pavlov/Stern Judging* and the *Contrite TFT/Simple Standing* pairs? In IR-context the key is the relationship to the second-order punishment. *Stern Judging* punishes agents (by giving *Bad* reputation) who cooperate i.e. who do not punish the *Bad* partner. Can we say that in IR-context that *Ego* who defected in the previous round (so (s)he was wrong) should punish *Alter* who cooperated previously? I have to accept it seems to be not very intuitive. But the other possible explanation also has a weak point. According to the widely accepted interpretation the *Pavlov* (*Win-Stay-Lose-Shift*) rule is a learning strategy that always pursues the maximal benefit. The *Pavlov* rule prescribes that after a *Temptation* situation when *Ego* defected and the *Alter* cooperated (so *Ego* totally "exploited" *Alter*), *Ego* should continue defection.

It is not easy to explain how the *Pavlov* rule can be the most successful cooperation providing strategy with this very egoistic element. This moment is crucial in the functioning of the *Pavlov* rule. When two *Pavlovian* agents play with each other and because of an error one of them defects instead of cooperating after a *DC* situation then the *Pavlov* rule prescribes mutual defection and in the next round both *Pavlov* players immediately react with *cooperation* to this *DD* situation. From this point of view this "exploitative" moment has a similar catalyzing character as the second-order punishment that the *Stern Judging* rule has in an IR-context. That is why it seems to be acceptable to state that this strategy pair can be characterized by the prescription: punish those who abstain from punishment. In his evolutionary approach to norms Axelrod (1986) also stressed the importance of the second-order punishment in the maintenance of norms.

The common feature of *Grim Trigger* and *Shunning* rule is that they easily lead to the end of the cooperation. They "condemn" all forms of action, no matter what *Ego* did in a *Bad* context (when *Alter* was *Bad* or *Ago* defected in the previous round), they prescribe defection in DR or *Bad* reputation in IR. They punish both punishment and abstaining from punishment.

The four principles which we can choose are the following: 'punish punishment', 'do not punish punishment', 'punish the non-punishment', and 'punish everything'. We can draw a moral lesson that we are what we choose, but it is not really important from a sociological point of view. These principles characterize community sentiments, so our analysis has to focus on how these principles work in communities. One principle may be dominant in a community at a given time, but it can be easily replaced with another one. The community's history can be characterized by a continuous wave that proceeds from the milder punishment "policy" to the increasingly stronger principles, and at the extreme point it turns back. The two "endpoint" are the *Simple TFT/Image Scoring* pair on one hand, and the *Grim Trigger/Shunning* strategy pair on the other hand. The former leads to the hypocrisy of cooperation at all costs, the latter to the punishment at all costs. Among them there are two moderate punishment policy principle pairs (*Contrite TFT/Simple Standing* and *Pavlov/Stern Judging*) which are the most successful at maintaining long-term cooperation. The simple fact that the argument of "deserved punishment" is incorporated into many narratives from the first children's tales to the Hollywood blockbusters means that the *Simple Standing* principle is important for every society. In "peaceful" times this principle is sufficient to maintain an adequate level of cooperation. When the community feels it is not enough, then the expectation for punishment increases, which is embodied in the criticism of abstaining from punishment (which is consistent with the *Stern Judging* principle). While these expectations are not widespread, the survival of cooperation is not threatened, but beyond a certain point the segregation (*Shunning*) logic becomes dominant. And when the people discover that this logic does not solve their cooperation problems, the pendulum starts in the opposite direction. The communities begin to relax the rigor of their punishment policy. Of course, this kind of social dynamic characterizes not so much society as a whole, but rather the smaller and larger communities and subcultures, and at any given moment these principles coexist with a wide range of movement, diversity, and intensities. Exploring these dynamics would require sensitive fieldwork in sociology and further simulation possibilities in evolutionary modeling.

## Acknowledgment

Support provided by the research project K83887 of the Hungarian Scientific Research Fund is gratefully acknowledged. I

would also like to thank Gábor Péli and the anonymous reviewer for reading and improving this paper.

## References

- Alexander, R.D., 1987. *The Biology of Moral Systems*. Aldine de Gruyter.
- Axelrod, R., 1984. *The Evolution of Cooperation*. Basic Books.
- Axelrod, R., 1986. An evolutionary approach to norms. *Am. Polit. Sci. Rev.* 80 (4), 1095–1111.
- Belnap, N., Perloff, M., Xu, M., 2001. *Facing the Future*. Oxford Univ. Press.
- Boerlijst, M.C., Nowak, M.A., Sigmund, K., 1997. The logic of contrition. *J. Theor. Biol.* 185, 281–293.
- Boyd, R., 1989. Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *J. Theor. Biol.* 136, 47–56.
- Boyd, R., Richerson, P.J., 1989. The evolution of indirect reciprocity. *Soc. Netw.* 11, 213–236.
- Boyd, R., Richerson, P.J., 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* 13, 171–195.
- Brandt, H., Sigmund, K., 2004. The logic of reprobation: assessment and action rules for indirect reciprocation. *J. Theor. Biol.* 231, 475–486.
- Frean, M.R., 1994. The Prisoner's dilemma without synchrony. *Proc. R. Soc. B* 257, 75–79.
- Friedman, J.W., 1971. A Non-cooperative Equilibrium for Supergames. *Rev. Econ. Stud.* 38 (1), 1–12.
- Hardin, G., 1982. Discriminating altruism. *Zygon* 17, 163–186.
- Hauert, Ch., Schuster, H.G., 1998. Extending the iterated Prisoner's Dilemma without synchrony. *J. Theor. Biol.* 192, 155–166.
- Imhofa, L.A., Fudenberg, D., Nowak, M.A., 2007. Tit-for-tat or win-stay, lose-shift? *J. Theor. Biol.* 247, 574–580.
- Kandori, M., 1992. Social norms and community enforcement. *Rev. Econ. Stud.* 59 (1), 63–80.
- Kanger, S., Kanger, H., 1966. Rights and parliamentarism. *Theoria* 32, 85–115.
- Kraines, D., Kraines, V., 1989. Pavlov and the Prisoner's Dilemma. *Theory Decis.* 26, 47–79.
- Kraines, D., Kraines, V., 1993. Learning to cooperate with Pavlov an adaptive strategy for the iterated Prisoner's Dilemma with noise. *Theory Decis.* 35 (2), 107–150.
- Malinowski, B., 1950. *Argonauts of the Western Pacific*. Routledge and Kegan Paul.
- Mauss, M., 1990. *The Gift: The Form and Reason for Exchange in Archaic Societies*. W W Norton and Company.
- Nowak, M.A., 2006. Evolutionary dynamics of cooperation. In: *Proceedings of the International Congress of Mathematicians*, pp. 1523–1540.
- Nowak, M.A., Sigmund, K., 1993. A strategy of win-stay lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 364, 56–58.
- Nowak, M.A., Sigmund, K., 1998. The dynamics of indirect reciprocity. *J. Theor. Biol.* 194, 561–574.
- Nowak, M.A., Sigmund, K., 2005. Evolution of indirect reciprocity. *Nature* 437, 1291–1298.
- Ohtsuki, H., Iwasa, Y., 2004. How should we define goodness?—reputation dynamics in indirect reciprocity. *J. Theor. Biol.* 231 (1), 107–120.
- Ohtsuki, H., Iwasa, Y., 2006. The leading eight: social norms that can maintain cooperation by indirect reciprocity. *J. Theor. Biol.* 239 (4), 435–444.
- Ohtsuki, H., Iwasa, Y., 2007. Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. *J. Theor. Biol.* 244, 518–531.
- Oliver, P., 1980. Rewards and punishments as selective incentives for collective action: theoretical investigations. *Am. J. Sociol.* 85 (6), 1356–1375.
- Pacheco, J.M., Santos, F.C., Chalub, F.A.C.C., 2006. Stern judging: a simple, successful norm which promotes cooperation under indirect reciprocity. *PLoS Comput. Biol.* 2 (12), 1634–1638.
- Panchanathan, K., Boyd, R., 2003. A tale of two defectors: the importance of standing for evolution of indirect reciprocity. *J. Theor. Biol.* 224, 115–126.
- Panchanathan, K., Boyd, R., 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432, 499–5024.
- Ridley, M., 1997. *The Origins of Virtue*. Softback Preview.
- Rosas, A., 2010. Evolutionary game theory meets social science: is there a unifying rule. *J. Theor. Biol.* 264, 450–456.
- Sahlins, M., 1972. *Stone Age Economics*. Aldine-Atherton.
- Searle, J., 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge Univ. Press.
- Sigmund, K., 2010. *The Calculus of Selfishness*. Princeton University Press.
- Sigmund, K., Nowak, M.A., 1998. Evolution of indirect reciprocity by image scoring. *Nature* 393, 573–577.
- Sugden, R., 1986. *The Economics of Rights, Co-operation and Welfare*. Palgrave MacMillan.
- Takahashi, N., Mashima, R., 2003. The Emergence of Indirect Reciprocity: Is the Standing Strategy the Answer? Hokkaido Univ. Working Paper No. 29. (<http://lynx.let.hokudai.ac.jp/COE21/pdf/029.zip>).
- Trivers, R., 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.
- Von Wright, G.H., 1963. *Norm and Action*. Routledge and Kegan Paul.
- Weber, M., 1978. *Economy and Society*. University of California Press.