Analyse & Kritik 22/2000 (© Lucius & Lucius, Stuttgart) p. 3–18

Peter Abell/Diane Reyniers

Generalised Reciprocity and Reputation in the Theory of Cooperation: A Framework

Abstract: We study the Iterated Bilateral Reciprocity game in which the need for help arises randomly. Players are heterogeneous with respect to 'neediness' i.e. probability of needing help. We find bounds on the amount of heterogeneity which can be tolerated for cooperation (all players help when asked to help) to be sustainable in a collectivity. We introduce the notion of Generalised Reciprocity. Individuals make a costly first move to benefit another under the reasonable expectation that either the other or somebody else will reciprocate. We hope that these tentative attempts at extending Axelrod's seminal work on cooperation will inspire future efforts in the field of organisational culture and social theory more generally.

1. Introduction

Social Theory is not a happy thing to contemplate; indeed, it is increasingly difficult to evade the uncomfortable conclusion that it comprises a failed intellectual tradition. In any one of its many guises Social Theory has not provided a convincing framework for deductively deriving collective social outcomes by modelling complex social mechanisms, starting from clear (micro) assumptions about individual actors and their (often strategic) interactions. A handful of seminal books, however, stand out as promising us something better and amongst the foremost of these is Robert Axelrod's prize winning volume *The Evolution of Cooperation* (1984). It provides a glimpse into a world where actors, when blessed with due foresight, strategically interact and, in so doing, apparently achieve a non-hierarchical solution to the classical Hobbesian problem. The book and a number of associated articles (Axelrod/Dion 1988; Axelrod/Hamilton 1981) have, along with parallel and contemporary achievements in evolutionary biology and game theory (Weibull 1995; Bendor/Swistak 1997), established a potential niche in social theory which is intellectually serious and shows every sign of becoming cumulative, whilst demonstrating a fine interplay between analytical and empirical (often simulated) results. It is entirely appropriate, then, that we should at the same time salute Axelrod and express the hope that the work he and his collaborators have inspired will gain a firm foothold in main-stream social theory. Our modest essay is designed to help achieve this latter objective.

We will explore how the insights which are to be found in the *Evolution of Cooperation*—and perhaps more generally within evolutionary game theory itself—might be brought to bear upon the analysis of *generalised reciprocity* in social systems. Generalised reciprocity arises when actor 1 will now cooperate with actor 2 on the expectation that *either* actor 2 *or* somebody else (actor 3) will reciprocate later, when actor 1 needs it. As such, generalised reciprocity is an inter-temporarily distributed asset available to a social group and may be appropriately conceived as a form of social capital.

Before embarking upon the general analysis it is important to establish clearly what Axelrod's work has and has not established for one encounters both extravagant and ill-informed claims from those who have only half-digested his contribution. For example, Axelrod does not claim, as is often attributed to him, that tit-for-tat (TFT) is superior to all other strategies in repeated prisoners dilemma (PD) situations. Indeed it is not. We shall start therefore by briefly reviewing the theory of cooperation.

2. The Theory of Cooperation in PD Type Situations

The tournaments which are detailed in Axelrod's book allow for repeated play between pairs of strategies and high-light, in particular, the virtues of TFT in this context. The number of repetitions is either fixed in advance or indeterminate (i.e. there exist a probability that no future engagements will arise). It seems, however, now more appropriate to construe his results within the general framework of evolutionary game theory whereby pairs of individuals are drawn at random from a population, each with a pre-assigned strategy, and obliged to play a *two-person* PD (with imperfect but complete information) under standard assumptions.[1] There always exist a significant probability that any pair of individuals will meet again in the future and repeat the strategically identical sort of encounter.[2] Each individual has full recollection of how each opponent has behaved in previous encounters and can, if it is strategically appropriate, take cognisance of the past in formulating current behaviour. In the usual interpretation, however, third parties cannot

[1] In fact this set-up corresponds to Axelrod's third type of simulation (where changing proportions of strategies in a population are introduced). He calls this an ecological simulation. The standard assumptions are those of the ranking of the values of outcomes and that the reward for co-operation is greater than half the reward for free-riding and being free-ridden upon.

[2] This along with the assumption that the future is not too-heavily discounted gives actors an incentive to anticipate the significance of the future in current behaviour.

observe the pair-wise interactions of others, so individuals (as strategic hosts) draw information about others solely from their own interactions; there are no extra-interaction reputation effects.[3]

We can distinguish between, on the one hand, non-conditional strategies (like always cooperate or defect or some fixed probabilities of both) and, on the other hand, conditional strategies (like TFT or tit-for-two-tats TF2T). It is only the latter which require some recollection. Given the basic pay-off structure, a fitness matrix can easily be computed giving the pay-off for each strategy (contingent and non-contingent) when pitted against itself and each other strategy: match this with the distribution of strategies in the population and the assumption of random pairings then the expected payoff for any given strategy (contingent upon the probability distribution of strategies) can be calculated.

Let us now incorporate Axelrod's principle—"whatever is successful is likely to appear more often in the future" (1984, 169) in terms of a dynamic specification whereby, with repeated play, the frequency of more successful strategies increases (shall we say, by emulation) at the expense of those which are less successful. Furthermore, allow for the possibility that established populations of 'native' strategies may be invaded by 'mutants' (i.e. previously unused strategies) then we may question how this will effect the dynamics.

Axelrod introduced what he termed 'collective stability' which is formally a Nash equilibrium (best reply) in which players use identical strategies. Furthermore, it is easy to show that TFT has this property (as long as the future features sufficiently) and this analytical result, along with the impressive results of the various tournaments, has promoted the view whereby TFT is in some way privileged and the key to solving the Hobbesian problem. For those blessed with sufficient foresight a cooperative equilibrium can apparently be attained without any appeal to other regarding sentiments like altruism. If everybody, in pairs, (or perhaps in n-collectives) plays TFT then we have a dynamically stable (N-person) Nash equilibrium. This is the interpretation commonly given to Axelrod's findings by those sociologists seeking to incorporate his ideas into their own work. However, such easy conclusions cannot be drawn (Bendor/Swistak 1997). What is it about TFT which is so special? First, we might note that the property of collective stability is only a necessary, not a sufficient condition for dynamic equilibrium. The issue then is, what are the *other* necessary and the *sufficient* properties of TFT which confer evolutionary advantage upon it (and indeed on similarly endowed strategies). Axelrod pointed to TFT's 'niceness', 'provocability' and 'forgivingness' and

[3] A natural extension would be to allow individuals to meet in an n-person PD (n>2). In general one would expect that as n increases, establishing co-operation would be more problematic. Axelrod 1988 reviews some of the work on n-person PD. The general problem is, of course, that if single defectors do not seriously erode collective co-operation then it may not be in the interest of the collective to punish the free-riding but to tolerate it.

minimal clustering property. TFT never defects first, responds to previous defection by defection and reverts to cooperation when faced with previous cooperation and can invade a population of 'all defect' if it has favoured interactions with itself. But, once again, the story is not as straightforward as this interpretation might imply.

We now know, courtesy of Boyd and Lorberbaum (1987) and Lorberbaum (1994), that no pure or mixed strategy is strongly stable in the sense that it can drive out all combinations of mutants. However, strategies can exhibit weak stability in the sense that they may consolidate at a certain frequency which cannot be driven down by any invading mutants (individual or jointly) though the mutants will remain as 'neutral' participants. Indeed, Bendor and Swistak (1995) establish an 'Evolutionary Folk theorem' for iterated PD which, not surprisingly, parallels the better known theorem for the repeated PD stage game. Accordingly, any level of cooperation in an evolutionary IPD is possible (given a sufficiently high likelihood of repetition). This being the case it cannot simply be niceness nor provocability which gives TFT any distinctiveness it possesses. Bendor and Swistak also discount clustering as the appropriate characteristic—a particularly significant negative finding for social theorists for the clustering of interactions is implied by most conceptions of social structure.

Axelrod believed that one of the characteristics of TFT which gives it superiority is its ability to invade an ecology of all Ds if TFT is minimally more likely to interact with itself than will all D. But once again it turns out that a wide variety of strategies have this propensity.

Bendor and Swistak argue that the key issue about the stability of a strategy is its minimal stabilising frequency (i.e. the minimal frequency which can ensure weak stability). No strategies (with sufficiently large likelihood of repetition) have a stabilising frequency below 50% but 'nice' and 'provocable' strategies become sufficient for weak stability at 50%, as the likelihood of repetition converges to unity.[4] This, however, is only a sufficient condition for weak stability and these authors introduce the idea of 'almost nice and retaliatory' to find the necessary conditions.

3. Intertemporal Bilateral Reciprocity

The standard theory of cooperation in IPD type situations which was set in motion so wonderfully by Robert Axelrod may provide a framework for analysing more complicated but realistic aspects of social systems many of which would extend the perspective of simultaneous to both sequential moves

[4] This result is only true under what is called proportional fitness rule dynamics (replicator dynamics). See Bendor/Swistak 1997.

and to reputation effects. I will scratch your back *now* if I have an expectation that you will scratch mine *when* needs be and I will scratch your back hoping others will observe and will be moved to scratch mine when appropriate. Do the characteristics of TFT extend naturally in this direction? Axelrod and Hamilton (1984) wrote:

> "...while the model treats the choices as simultaneous, it would make little difference if they were treated as sequential. We wonder if this is correct?"

Actor 1 makes himself vulnerable by moving first, expecting (hoping) that the recipient of her attention will reciprocate when necessary. It is the intertemporal sequencing of the (supposedly personally costly) activities which generates a PD type game with perfect information. Actors must, of course, have a 'significant' likelihood of meeting again to occasion possible reciprocation and also not discount the future too heavily. Furthermore, these ideas must extend to *repeated* intertemporal encounters where either actor may on differing occasions assume the role of first mover. Indeed, in many social systems this sort of individualised bilateral reciprocity appears to develop rather painlessly and people may keep a rough and ready running tally of their respective levels of outstanding credit or debt against a background of previously successfully accomplished exchange. Norms arise (which incidentally are little understood) which render fungible the relative values of different sorts of social exchange and which probably prevent over-indebtedness. A sense of justice or injustice often emerges in the context of the equitability of exchanges (Binmore 1997) and though it is beyond the scope of this paper, it is the sense which can lead to altruistic (i.e. other regarding) sentiments (Abell 1996).

These latter considerations apart, norms of reciprocity arise and consequently a role structure whereby all are subject to normative expectations that (a) when appropriate they will make the first move (e.g. 'help someone') and (b) when they themselves have been helped they will reciprocate. It should be emphasised that though such patterns may give the impression of altruistic behaviour (indeed, in some places they have been labelled reciprocal altruism) no assumption of other-regard is involved. The exchange arises from prudent self-interest. Axelrod's tournaments are constructed in a manner whereby the (role) distinction between first-mover and reciprocator does not apply, because the underlying PD is played as a game of imperfect information (i.e. coincidental moves). Does this difference have any significance?

Consider a pair of players and assume to whom the role befalls to be first-mover in the course of their repeated encounters is random (or equivalently the need for help arises randomly) then situations can arise where it is the same player who is repeatedly asked to cooperate (i.e. lend help) without any, as yet, call for reciprocation on the other's behalf. Thus, purely by chance credits and debits can accumulate (in the short term) in a one-sided manner. The situation arises from the game depicted in Fig Ia which reduces to a PD if the actors are by chance selected in sequence.

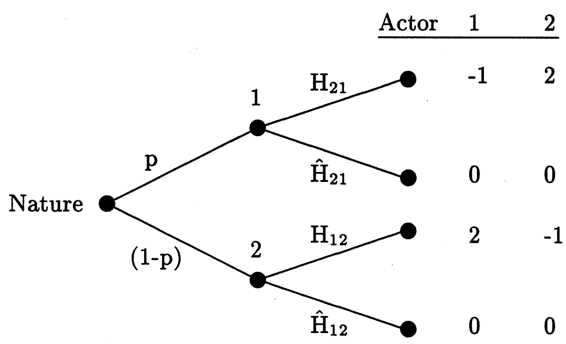


Figure Ia: Selection of first mover

The model is certainly nearer to reality than the immediate reciprocation model of IPD. What may be required now is what we might term TIT for accumulated TAT (TFAT). The first-mover needs to be repeatedly 'nice' and 'patient' until the time arises when she is in need of reciprocation.

If we were to allow reputation effects beyond any pair of interactants then the first-mover's patience may be dependent upon how the recipient has behaved in interaction with others. If the recipient has indicated a willingness to co-operate with third parties then this could encourage patience if the random need for help requires it. Mechanisms of this sort could be incorporated into simulations as could the social structures of who observes whom if observability is not complete (Nowak/Sigmund 1998).

Over time, of course, large numbers of interactions will guarantee that the reciprocation will even out, if both individuals pursue TFAT. One's intuition is that the analytical findings in respect of TFT should all go through though they will be more sensitive to the 'shadow of the future'. However, since isolated pairs of interactants are relatively rare in real social systems we move onto similar issues in the context of generalised reciprocity.[5]

In the game depicted in Figure Ia exactly one player needs help in each period. The Bilateral Reciprocity (BR) game of Figure Ib models the perhaps more realistic scenario where nobody or one or both players may need help.

[5] Currently plans are afoot to simulate this sort of situation with a varying patience parameter in TFAT.

Player 1 and player 2 need help with probability p_1 and p_2 respectively. If one and only one player needs help, the other player decides whether to help. If both players need help, they decide simultaneously whether to offer help. The standard PD is embedded in BR as the imperfect information subgame in which both players need help. Whenever a player gets help he receives a benefit $b_i (i = 1, 2)$ and whenever he offers help he incurs a cost $c_i (i = 1, 2)$. In the iterated version of the game (IBR) there is a probability δ in any period that the game will continue for a further period, or, equivalently, the game is repeated with infinite horizon and δ is the discount factor. The question we want to address is whether there is a Nash equilibrium in which both players always offer help.

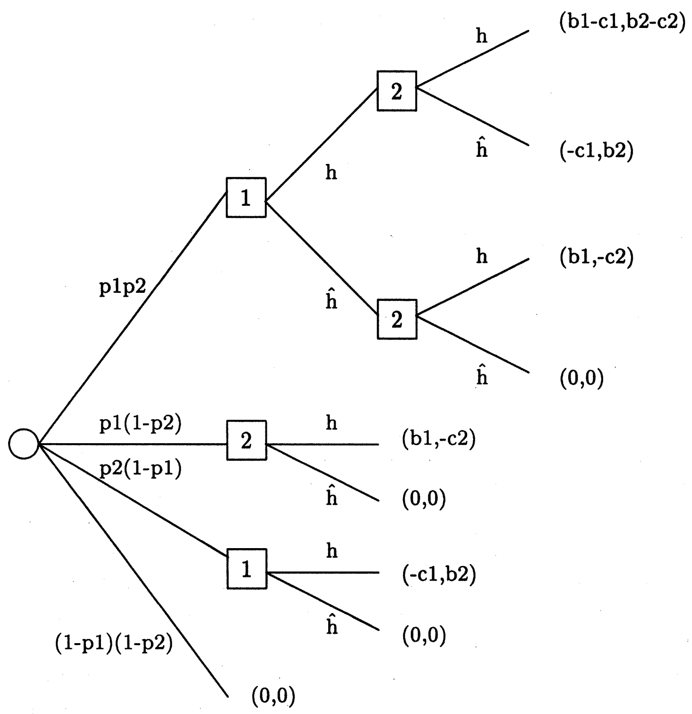


Figure Ib: Bilateral reciprocity game

The equivalent of the 'grim strategy' in IPD in this game is for a player to always offer help until the other player refuses to help and to refuse to help from then onwards. Of course, if both players playing the grim strategy is a **NE**, then cooperation can be sustained i.e. help is always offered when needed. So assume player 1 plays the grim strategy, what is player 2's best response? If player 2 also adopts the grim strategy, his expected payoff is

$$(p_2 b_2 - p_1 c_2)\frac{1}{1-\delta}. \tag{1}$$

Since he will get help with probability p_2 and will offer help with probability p_1. If player 2 decides to deviate, i.e. not offer help when player 1 needs it, he will still receive help from player 1. In IBR a player does not detect deviation until he needs help and is denied. A deviating player 2 (who never gives help) therefore receives an expected payoff of

$$p_2 b_2 + \delta p_2 b_2 (1-p_1) + \delta^2 p_2 b_2 (1-p_1)^2 + \ldots = \frac{p_2 b_2}{1-\delta(1-p_1)}. \tag{2}$$

Hence, the grim strategy is a best response if (1) exceeds (2) or

$$\delta\left(p_2\frac{b_2}{c_2} + 1 - p_1\right) \geq 1.$$

Thus we conclude that cooperation can be sustained in IBR if

$$\delta\left(p_2\frac{b_2}{c_2} + 1 - p_1\right) \geq 1 \text{ and } \delta\left(p_1\frac{b_1}{c_1} + 1 - p_2\right) \geq 1.$$

If the players are identical, these conditions reduce to

$$\delta\left(1 + p\left(\frac{b}{c} - 1\right)\right) \geq 1. \tag{3}$$

Note that we need $b > c$ for identical players to support cooperation.

The more often we need each other (p large), the longer the expected duration of the relationship or the more patient we are (δ large), the larger the benefit we receive when helped (b large) and the lower the cost of giving help (c small), the more likely it is we are able to sustain cooperation.

From (3) we can also conclude that it is harder to obtain cooperation in IBR than in IPD since IBR reduces to IPD for $p = 1$.

Let us now turn our attention to an n-player game of bilateral reciprocity. In each period player i needs help with probability p_i. A player who needs help randomly asks one of the other $n-1$ players for help with the proviso that each player can be asked to help by at most one other player. In other respects the setup is as in IBR. In particular, the grim strategy involves ceasing cooperation as soon as any player deviates. Player i is then asked to help with probability $(\sum_{j\neq i} p_j)/(n-1)$. Repeating the analysis for 2 player IBR leads to the result that cooperation can be sustained if

$$\delta\left(p_i\frac{b_i}{c_i} + 1 - \frac{\sum_{j\neq i} p_j}{n-1}\right) \geq 1 \text{ for all } i. \tag{4}$$

Note that if players are identical (4) reduces to (3). This implies that there is no group size effect for identical players. This is not suprising since we are dealing with bilateral interaction. While the number of people needing help increases with n, so does the number of people available to help. For heterogeneous players however, (4) can be interpreted as a constraint on the amount of heterogeneity which can be tolerated for cooperation to be sustainable in a collectivity. Heterogeneity is more tolerable the longer the expected duration of the game. Here we find a group size effect since heterogeneity is likely to increase as a group gets larger. (We could, for example, envisage a model where the p_i's are drawn from a given distribution).

Our analysis of n-player IBR can shed some light on the prospects for cooperation when internally homogeneous collectivities are merged. Consider a special case of IBR with two types of players: n_1 'needy' and n_2 'self-sufficient' players ($n_1 + n_2 = n$) have probability p_1 and p_2 respectively of needing help ($p_1 > p_2$). Further assume that everyone's benefit to cost ratio equals $\frac{b}{c}$. For large n_1 and n_2 condition (4) reduces to

$$\frac{n_2}{n} \geq \frac{\frac{1}{\delta} - 1 + p_1 - p_2(\frac{b}{c})}{p_1 - p_2}. \tag{5}$$

Clearly, (5) is impossible to satisfy if the right hand side exceeds 1, i.e. when $p_2(\frac{b}{c} - 1) < \frac{1}{\delta} - 1$. The implication of this result is that if cooperation cannot be sustained in the self-sufficient group then it cannot be sustained in the merged group.

The question then arises as to whether we need cooperation to be sustainable in the needy group to get cooperation in the merged group. To answer this question, rewrite (5) as

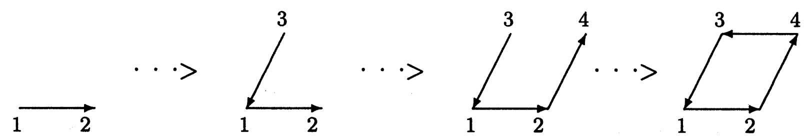
$$\left(\frac{b}{c} - \frac{n_2}{n}\right) p_2 \geq \frac{1}{\delta} - 1 + \left(1 - \frac{n_2}{n}\right) p_1. \tag{6}$$

If cooperation cannot be sustained in the needy group, i.e. $\frac{1}{\delta} - 1 > p_1(\frac{b}{c} - 1)$, then the right hand side in (6) exceeds $p_1(\frac{b}{c} - \frac{n_2}{n})$ and hence (6) can never be satisfied. So, if cooperation cannot be sustained in the needy group then it cannot be sustained in the merged group. Only if cooperation is sustainable within both groups is it possible to sustain it in the merged group. Even when it is possible to sustain cooperation in the merged group, the self-sufficient group is obviously better off if it remains independent. If we think of 'neediness' as a characteristic of corporate culture, these results help us understand the difficulties of merging corporate cultures.

An interesting and more realistic version of IBR is the game in which p_i is private information. Players could then learn about each other's neediness through repeated interaction. In this scenario it may be optimal for a player with high neediness to hide this fact and not always ask for help when he needs it since asking for help increases other player's estimate of his neediness and they may stop cooperating. Staying in the collectivity where his needs are not always satisfied may be better than not staying. However, since we have not analysed this incomplete information game, these are just conjectures which might be studied in further work.

4. Generalised Reciprocity and Reputations

Reliance upon bilateral reciprocity does not ultimately confer efficiency upon social organisations. Most social systems—particularly those of any size—come to depend not upon bilateral but on generalised reciprocity. The transition from patterns of bilateral to generalised reciprocity is one that is most important to the durability of any social collectivity. It enables individuals (actor-1) to prudently make the first (costly) move to benefit another (actor-2) under the reasonable expectation that *either* actor 2 *or* somebody else (actor 3) will reciprocate on the occasion when actor-1 requires it. That expectations, so formulated, can be in alternative form, confers efficiency advantages in the sense that when actor-1 requires reciprocation (e.g. help) it may be that actor-2 (who is in debt to actor-1) will not be available, (assume availability is randomly distributed) though somebody else (actor-3) could be. It will, of course, be appropriate that, on yet another occasion, actor-2 will reciprocate in respect of either actor-3 or somebody else who has 'helped' actor-3. It may prove useful to have a diagram to keep track of these sorts of mechanisms. In Fig I an emerging 'cooperative cycle' is illustrated whereby 'help debts' are eventually settled around a 4-cycle (Boyd/Richerson 1989).



Key: i $\rightarrow$ j   i 'helps' j

Figure I: Generalised reciprocity

Clearly, rather more ambitiously, we could conceive of a time dependent (actor x actor) matrix whose entries would represent, at a particular time, either the number of discrete cooperative acts (e.g. helping) or a continuous measure thereof. Then, the in and out-degree would give a portrayal of balance of debits and credits.

The strategic structure of Generalised reciprocity (in simple binary form) may be deemed to take one of three forms depicted in Figs II(a),(b) and (c) respectively. They vary in their information structure.

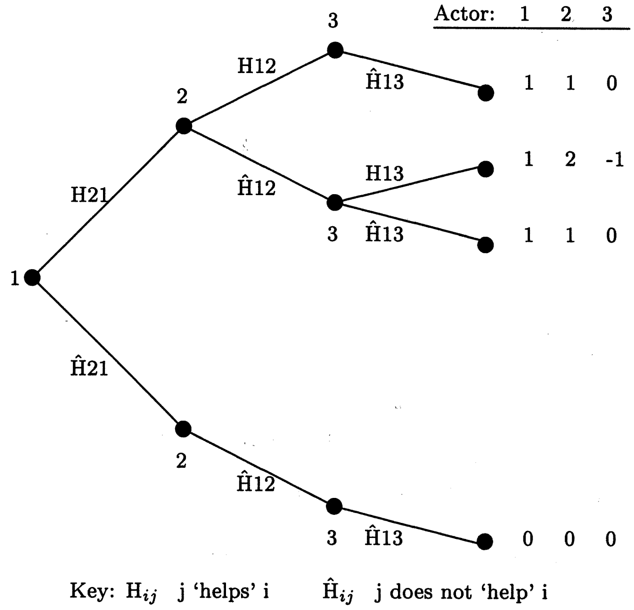


Figure IIa: Models of generalised reciprocity

In Fig II (a) ('all is visible') actor 1 has to decide whether or not to help actor 2 (who currently needs help). Everybody concerned (actor 2 and 3) observe 1's decision. At a subsequent juncture actor 2 has to decide whether or not to reciprocate in respect of actor 1. Actor 3 observes what actor 2 does. If actor 2 does reciprocate then actor 3 has an easy decision—there is no need for *generalised reciprocity.* If, however, actor 3 observes that actor 2 does not reciprocate she has herself to decide whether or not to do so. The Nash Equilibrium is, of course, (by backward induction) the all no help (0,0,0) outcome in a one-shot situation. Note also that, in the context of reputations, actor 3 is not merely a first mover as she has information about actor 1's willingness to help. If we combine third party reputation effects with generalised reciprocity then actors may be moved to help those with reputations for helping, estimating an appropriate probability of reciprocity, but also giving an eye to their own reputations and, thus inclined to offer help.

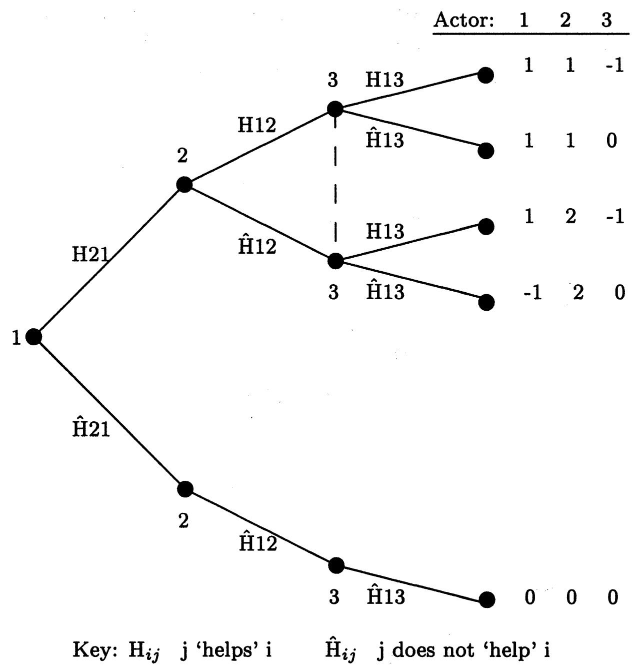
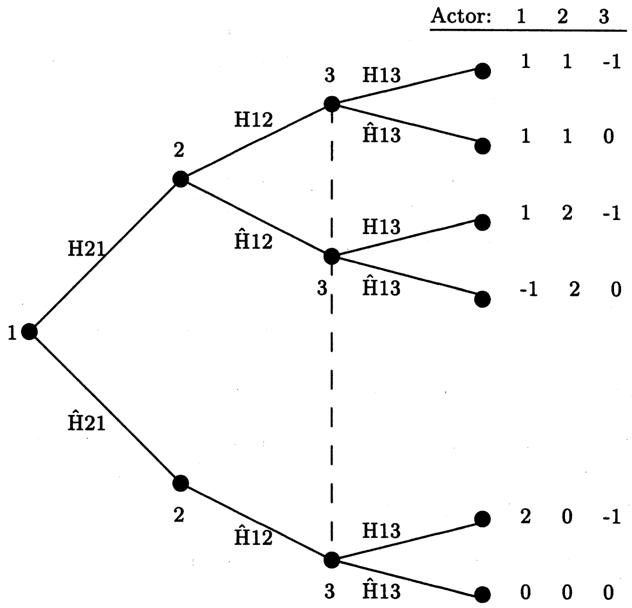


Figure IIb: Models of generalised reciprocity

Fig II(b) is slightly different: again actor 1 has to decide whether to make the first move. Both actor 1 and 2 observe this move. Subsequently actor 2 needs to decide whether or not to reciprocate. However, now neither actor 2 nor 3 knows whether the other will help. The one shot Nash equilibrium is again (0,0,0).

In Fig II(c) actor 1 makes the first move, in respect of actor 2 and actor 2 has to decide whether to reciprocate. 3 now, however, does not know whether 1 helped 2 nor whether 2 will help 1 and has to decide whether to help! Again the one-shot Nash equilibrium is at (0,0,0).



Key: H_{ij} j 'helps' i $\hat{H}_{ij}$ j does not 'help' i

Figure IIc: Models of generalised reciprocity

We shall concentrate solely upon the strategic situation depicted in Fig II(a) (all is visible) and imagine a context where the game is to be repeated with a sufficient probability for the shadow of the future to obtain between individuals drawn from a large population.

The simple (non-contingent) strategies available to the three actors are as follows:

Strategies for Actor 1:

S_{11}	help actor 2	H_{21}
S_{21}	do not help actor 2	$\bar{H}_{21}$

Strategies for Actor 2:

S_{12}	(If H_{21} then H_{12}) and (if $\bar{H}_{21}$ then $\bar{H}_{12}$)
S_{22}	(If H_{21} then $\bar{H}_{12}$) and (if $\bar{H}_{21}$ then $\bar{H}_{12}$)

Strategies for Actor 3:

$$S_{13}\begin{cases}(\text{If } H_{21} \text{ and } H_{12} \text{ then } \bar{H}_{13}\\(\text{If } H_{21} \text{ and } \bar{H}_{12} \text{ then } H_{13}\\(\text{If } \bar{H}_{21} \text{ and } \bar{H}_{12} \text{ then } \bar{H}_{13}\end{cases}$$

$$S_{23}\begin{cases}(\text{If } H_{21} \text{ and } H_{12} \text{ then } \bar{H}_{13}\\(\text{If } H_{21} \text{ and } \bar{H}_{12} \text{ then } \bar{H}_{13}\\(\text{If } \bar{H}_{21} \text{ and } \bar{H}_{12} \text{ then } \bar{H}_{13}\end{cases}$$

If we now introduce the idea of 'availability' of an actor to 'give help', the sequential structure may be envisaged in a number of ways. Consider:

(i). Actor 2 needs help (2 is selected at random Fig I(a)).

(ii). Actor 1 is selected at random as a candidate to help actor 2. Actors 1 and 2 are both aware of the outcome of any past interactions between them (either as helper or helpee).

(iii). Actor 1 is either ***available or not available*** (with a given probability) to help actor 2.

(iv). If actor 1 is available then she helps or does not help (as a constituent move in a strategy (Raub/Weesie 1990)).

(v). If actor 1 is not available or chooses not to help then actor 3 is selected at random to help actor 2 (actor 2 is removed from the selection).

(vi). If 3 is not available or does not help then the sequence may be iterated without replacement for x other actors ($x \leq N-2$) search depth). In pursuit of realism this basic picture may be modified in a number of additional respects.

(vii). Actor 2 may *select* others as helpers on the basis of past interactions. The potential field might be *all* others or a neighbourhood only.

(viii). Allow for reputation effects so actors can observe (some or all) interactions between others and formulate strategies appropriately.

(ix). Allow for parallel need for help, so there may be competition for help (thus, availability). Actors can only help one (or more?) persons at a time.

(x). Allow for variations in depth of search.

(xi). Allow for a "broadcast" of the need for help across $(N-1)$ others.

An alternative way to picture the situation is as follows: consider a population of N individuals each of which may, at some juncture, play any one of the roles of actor 1,2 and 3. These roles may be described as the role set. Start by assuming that every individual in the population is related to every other individual so that each pair of individuals can assume either of the three pairs of strategically related roles.

Next, assume that triples of individuals are selected at random (or the selection may be envisaged as driven by a random occurrence for the 'need for help'), then further assume the individuals are also randomly assigned to one of each of the three roles. Thus, any individual may adopt any one of eight possible strategies across the three roles as (if needed for clarity) in Fig III. For example, (S_{11}, S_{12}, S_{13}) implies that if the individual is selected to play the role of actor 1 she will play S_{11}, if the role of actor 2, S_{12} and, finally, if the role of actor 3, S_{13}.

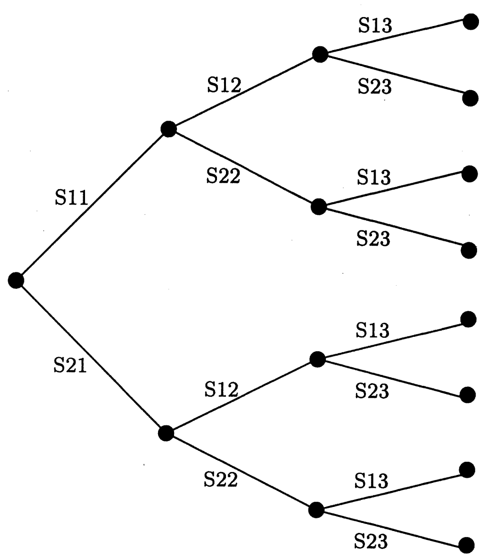


Figure III: Possible strategy combinations

Whichever way we are inclined to view generalised reciprocity and reputations the question arises as to what the equivalent of TFT might be—if, indeed, there is one—in these more complex situations. Can cooperation be achieved and, if so, more or less easily than in the simpler situations analysed by Axelrod. We hope future analytical work and simulations will provide an answer.

Bibliography

Abell, P. (1996), A Model of the Informal Structure (Culture) of Organisations, in: *Rationality and Society 8*, 433–452

Axelrod, R. (1984), The Evolution of Cooperation, New York.

— /Dion, D. (1988), The Further Evolution of Cooperation, in: *Science 242*, 1385–90

— /Hamilton, N. (1981), The Evolution of Cooperation, in: *Science 211*, 1390–1396

Bendor, J./P. Swistak (1995), Types of Evaluationary Stability and the Problem of Cooperation, in: *Proceedings of the National Academy of Science, USA 92*, 3596–3600

— / — (1997), The Evolutionary Stability of Cooperation, in: *American Political Science Review 91*, No 2, 290–307

Binmore, K. (1997), *Just Playing: Game Theory and the Social Contract II*, Cambridge

Boyd, R./J. Lorberbaum (1987), No Pure Strategy is Evolutionary Stable in the Repeated Prisoners' Dilemma, in: *Nature 327*, 58–59

— /P. J. Richerson (1989), The Evolution of Indirect Reciprocity, in: *Social Networks II*, 213–236

Lorberbaum, J. (1994), No Strategy is Evolutionary Stable in Repeated Prisoners' Dilemma, in: *Journal of Theoretical Biology 168*, 117-130

Nowak, A./K. Sigmund (1998), Evolution of Indirect Reciprocity by Image Scoring, in: *Nature 393*, 573–577

Raub, W./J. Weesie (1990), Reputation and Efficiency in Social Interactions: An Example of Network Effects, in: *American Journal of Sociology 96*, 626–54

Weibull, J. (1995), *Evolutionary Game Theory*, Cambridge