ECAI 2016
G.A. Kaminka et al. (Eds.)

doi:10.3233/978-1-61499-672-9-417

The Game of Reciprocation Habits

Gleb Polevoy and **Mathijs de Weerdt** and **Catholijn Jonker**[1]

Abstract. People often have reciprocal habits, almost automatically responding to others' actions. A robot who interacts with humans may also reciprocate, in order to come across natural and be predictable. We aim to facilitate decision support that advises on utility-efficient habits in these interactions. To this end, given a model for reciprocation behavior with parameters that represent habits, we define a game that describes what habit one should adopt to increase the utility of the process. This paper concentrates on two agents. The used model defines that an agent's action is a weighted combination of the other's previous actions (reacting) and either i) her innate kindness, or ii) her own previous action (inertia). In order to analyze what happens when everyone reciprocates rationally, we define a game where an agent may choose her habit, which is either her reciprocation attitude (i or ii), or both her reciprocation attitude and weight. We characterize the Nash equilibria of these games and consider their efficiency. We find that the less kind agents should adjust to the kinder agents to improve both their own utility as well as the social welfare. This constitutes advice on improving cooperation and explains real life phenomena in human interaction, such as the societal benefits from adopting the behavior of the kindest person, or becoming more polite as one grows up.

1 Introduction

Interaction is central in human behavior, e.g., at school, in file sharing over networks, and in business cooperation. While interacting, people tend to reciprocate, i.e., react on the past actions of others [9, 11, 14]. Imagine software agents owned by individuals repeatedly competing with the same people online. People expect reciprocal behavior and tend to behave so themselves. Virtual assistants also need to be reciprocal in order to be credible. Countries at an arms race or arguing friends also tend to be nicer if the other side is nicer [7, 27, 13]. In these and other cases of repeated interaction, we can help people and artificial agents obtain more from the interaction by providing decision support. The decision is how to reciprocate. Reciprocating efficiently includes defining to one's software agent or other artificial agents how to reciprocate with humans. In order to help people strategically choose efficient approaches for reciprocating, and to predict that strategic choice of how to reciprocate, a model is needed that is amenable to analytical analysis and has enough predictive power.

Consider the following example of an arms race.

Example 1 *Consider n countries $1, 2, \ldots, n$; each country can put a certain arsenal of weapons at the border with its neighbors. What a country approximately does with respect to another country at a given year is what was done in the previous year, adjusted to react to what the other countries did. If they armed themselves against us, we also will, and if the others aimed at us less, so shall we. This process is often reciprocal with linear reactions [7, 27]. Perhaps, one reason for that is that politicians can explain a reciprocal action as a proper reaction to the nation. A crucial question is how to make this process efficient, so that one's country, and, preferably, everyone incurs the least possible cost.*

In this example, an action had a negative influence on the other country. We can also consider a positive influence on the other side in this context; for instance, a concession.

Software agents can reciprocate automatically.

Example 2 *Consider software agents running on computers in a cloud. They need to agree on how much resources each is allocated. Since their owners may want to be nice to others reciprocally, it is reasonable to make them reciprocate. Everyone wants her agent to reciprocate as efficiently as possible, and also the society can save much money by efficient reciprocation.*

Companies can reciprocate while achieving mutual gain.

Example 3 *Reciprocation is useful in business life [25]. Reciprocating means helping the other, for example, by redirecting potential clients to another company. It is definitely economically important to make this reciprocation efficient.*

The existing studies of reciprocation (sometimes repeated) either attempt to explain why reciprocation is there in the first place [4, 3, 26, 10], or, given that reciprocation exists, they analyze what happens in a short interaction where being reciprocal pays off [5, 9, 23]. We, on the other hand, consider a lengthy interaction, that is (naturally) bound to be reciprocal, but changing the approach of reciprocation is possible, in order to receive more and do less.

To study such interactions, we employ the model from Polevoy, de Weerdt and Jonker [22][2], which formally defined and analyzed repeated intrinsic reciprocation, to understand how reciprocity makes interaction evolve with time. We briefly summarize the model. Actions, which are influences of an agent on another one, are represented by *weight*, where a higher value means a more desirable contribution to its recipient. That model was mainly inspired by arms race models [7, 27] and a model of spouses arguments [13]. Given the model, the paper [22] analyzes the interaction it engenders.

[1] Delft University of Technology, email: g.polevoy@tudelft.nl

[2] The full version can be found at `http://arxiv.org/abs/1601.07965`.

This model consists of two reciprocation attitudes, where the action of an agent is a convex combination[3] between i) one's own kindness or ii) one's own last action (mental inertia), and the other's last action (reaction). The combination is determined by the agent's reciprocation coefficient. Since the last own action is, recursively, a product of previous actions, it represents the agent at a given time, including her history. Attitude i), which is connected to kindness, is called *fixed*, and ii) depending on one's own last action is called *floating*.

A reciprocation process converges, and in many cases, the actions in the limit are found in [22]. We aim to to provide decision support and predict the strategic reciprocation. A natural question to ask here is in what way the agents can strategically influence the reciprocation process for their own good, and what the social welfare will become when every individual behaves strategically. Setting one's way of reciprocating resembles Mastenbroek's [17, Chapter 14] recommendation to know one's own negotiating style and adjust it. Assuming people strategically choose each action is unrealistic, since people usually act on habits [15], and a strategic choice consists of choosing a habit for the reciprocal interaction. Here, the habit, chosen after deliberation, can be the balance between reacting and being faithful to oneself, as defined in the model. It is also easy to prescribe a "habit" to a robot.

Choosing habits resembles bounded rationality, especially that of procedures of choice [24, Chapter 2]. Indeed, our agent follows the procedure of rationally choosing among the possible habits. The difference is that choosing a habit does include a rational step, and is, therefore, amenable to a standard game-theoretic analysis, like NE and price of anarchy and stability. Choosing habits resembles metagames as well, when an agent chooses a representative to play the underlying game for her. For instance, Rubinstein [24, Chapter 8] and [21, Chapter 9] define a *machine game*, where an agent wants a well-paying strategy that is simple to implement. This trade-off is modeled by choosing a finite deterministic automaton to play the repeated game, where the agent's utility increases in the utility of the underlying game and decreases in the number of the states of the chosen automaton. The equilibria in this game are found for the case of the utility of the repeated game being defined as the limit-of-means or with discounting in [6]. A player in a machine game chooses a finite automaton, while our player chooses a habit. Choosing an automaton, however, considers the bounding effect of finiteness and attempt to minimize the automaton's state space, while we simply consider a best possible habit, all habits being equally simple. Therefore, our model neither generalizes theirs nor is our model generalized by theirs. Additionally, no finite automaton is able to model reciprocation, though it is possible to approximate it arbitrarily.

To model strategically setting one's habits, we define the utility of an agent and then we consider the one-shot game of setting one's own reciprocation attitude or coefficient, each of which represents a habit. We analyze changing reciprocation attitude for a pairwise interaction. Pairwise interactions still allow for many agents provided assuming that the agents do not mix one relationship with the other ones.

All the agents choose their reciprocation habits and then the reciprocation process plays itself. Our contributions include a characterization of this game's Nash equilibria (NE) and a discussion of their efficiencies. We consider only pure NE in this paper. Analyzing this game provides an insight into how people and machines could change their behavior to achieve a more desirable behavior in the limit of the interaction process. This desirability can be to themselves or to the society. In addition to predicting the strategic reciprocation and advising on what to do, the analysis explains the following known phenomena. First, in reciprocation, we often notice that when the example of the kindest person is followed by others, it makes the group more successful [2]. We also notice, that people tend to become more polite as they grow up [12], which is yet another example of the utility of learning from the behavior of the kindest.

We present the model in Section 2. To make this paper self-contained, Section 2.3 provides the necessary background. We consider the game of choosing the reciprocation attitude in sections 3, 4 and 5, proving the central Theorems 4 and 5. We also model in Section 6 what happens if an agent can choose both own attitude and reciprocation coefficient. The answers are given in the key Theorems 6 and 7.

We briefly describe the model and the results for n agents in Section 7. We deal with convergence of the best response dynamics to a NE in Section 8 and conclude in Section 9.

2 Modeling Reciprocation

We first model agents, times and actions. We conclude the section by sharpening the model and providing explanatory examples. Let $N = \{1, 2\}$ be $n = 2$ interacting agents. Time is modeled by a set of discrete moments $t \in T \triangleq \{0, 1, 2, \ldots\}$[4], defining a time slot when the agents act.

Denote the weight of an action by agent $i \in N$ on another agent $j \in N$ at moment $t \in T$ by $x_{i,j} \colon T \to \mathbb{R}$. For example, when interacting by file sharing, the actions of sending a valid piece of a file, nothing, or a piece with a virus are decreasing in weight. Since only the weight of an action is relevant, we usually write "action" while referring to its weight.

We now define two reciprocation attitudes, which define how an agent reciprocates. We need the following notions. The kindness of agent i is denoted by $k_i \in \mathbb{R}$; w.l.o.g., $k_2 \geq k_1$ throughout the paper. Kindness models inherent inclination to help others; in particular, it determines the first action of an agent, before others have acted. We model agent i's inclination to mimic the other agent's action by reciprocation coefficients $r_i \in [0, 1]$. Here, r_i is the fraction of $x_{i,j}(t)$ that is determined by the last action of j upon i. Conceptually, reacting to last actions, one reacts to the actor, since "who you are is what you do" [18].

Intuitively, with the *fixed* attitude, actions always depend on the agent's kindness, while the *floating* attitude moves freely in the reciprocation process, and kindness directly influences such behavior only at $t = 0$. In both cases $x_{i,j}(0) \triangleq k_i$.

Definition 1 *For the* fixed *reciprocation attitude, agent i's action on another agent j is determined by j's last action weighted by r_i and by the agent's kindness weighted by $1 - r_i$.*

[3] A combination is convex if it has nonnegative weights that sum up to 1.

[4] Allowing agents to be non-synchronized is possible, but we assume synchroneity for the sake of clarity.

That is, for $t \in T$,

$$x_{i,j}(t) \triangleq (1 - r_i) \cdot k_i + r_i \cdot x_{j,i}(t-1).$$

Definition 2 *In the* floating reciprocation attitude*, agent i's action is a weighted average of that of the other agent j, and of her own last action. To be precise, for $t \in T$,*

$$x_{i,j}(t) \triangleq (1 - r_i) \cdot x_{i,j}(t-1) + r_i \cdot x_{j,i}(t-1).$$

The relations are (usually inhomogeneous) linear recurrences with constant coefficients, but many variables. We could express the dependence of $x_{i,j}(t)$ only on $x_{i,j}(t')$ with $t' < t$, but then the coefficients would not be constant, besides the case of two *fixed* agents. We are not aware of a method to use the general recurrence theory to improve our results.

2.1 Context and Examples

Compared to the other reciprocation models, our model takes reciprocal actions as given and looks at the process, while other models either consider how reciprocation originates, such as the evolutionary model of Axelrod [4], or take it as given and consider specific games, such as in [5, 8, 9, 23].

In Example 1, let the reciprocation coefficients be $r_1 = 0.2, r_2 = 0.7$. Assume the kindness to be $k_1 = 0$ and $k_2 = 0.5$. At $t = 0$, every country's action on every other country is equal to her kindness value, so $x_{i,2} = 0$ and $x_{2,1} = 0.5$. If all countries rely on their previous action, meaning that they are *floating*, then, at $t = 1$ they act as follows: $x_{1,2}(1) = (1 - 0.2) \cdot 0 + 0.2 \cdot 0.5 = 0.1$, $x_{2,1}(1) = (1 - 0.7) \cdot 0.5 + 0.7 \cdot 0 = 0.15$. Theorem 2 implies they converge to the common limit $\frac{0.7}{0.2+0.7} \cdot 0 + \frac{0.2}{0.2+0.7} \cdot 0.5 = 1/9$. This is closer to k_1 than to k_2, since country 1 is less responsive, in the sense that $r_1 < r_2$.

Consider modeling tit for tat [3]:

Example 4 *In our model, a tit for tat agent with two options: cooperate or defect is easily modeled with $r_i = 1$, $k_i = 1$, meaning that the original action is cooperating (1) and the next action is the current action of the other agent. If one of two tit-for-tat agents makes a mistake and begins with defection ($k_2 = 0$), then they will alternate.*

If the agents are human, this example predicts an indefinitely long alternation, which seems unrealistic to us. Similarly, an agent that sticks to his actions regardless the other seems highly implausible. This provides evidence that extreme values of the reciprocation coefficients are uncommon in life.

2.2 Utility Definition

An agent's utility at a given time moment is the action one receives minus the effort incurred by the action one performs. Colloquially, this is what the agent gets minus what she gives. This classical way of defining utility is expressed, for instance, in the quasilinear preferences of auction theory [20, Chapter 9.3]. Formally, we define as follows.

Definition 3 *The utility of agent i at moment t, $u_{i,t} \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, is defined as*

$$u_{i,t}\left(x_{i,j}(t), x_{j,i}(t)\right) \triangleq x_{j,i}(t) - \beta_i x_{i,j}(t),$$

where β_i is the relative importance of the effort incurred by performed actions for i's utility. The personal price of acting is higher, equal or lower than of receiving an action, if β is bigger, equal or smaller than 1, respectively.

Denote $x(t) \triangleq x_{1,2}(t)$ and $y(t) \triangleq x_{2,1}(t)$. Thus, at time t, agent 1's utility is $y(t) - \beta_1 x(t)$ and 2's utility is $x(t) - \beta_2 y(t)$.

We take acting with a minus sign, to account for the effort it takes (a negative β_i would mean that the agent enjoys making effort). According to this formula, when $\beta_i > 0$, a negative action would suddenly contribute to the utility; we needed to take the absolute value. Instead, we will assume that actions are always non-negative, which is equivalent to all kindness values being non-negative. We still can have negative influence, we have simply mathematically transformed all the original kindness values by adding a sufficiently large number so that they all have become nonnegative.

To model the utility in the long run, we give the following Define the asymptotic utility, or just the *utility*, of agent i, $u_i \colon \mathbb{R}^\infty \times \mathbb{R}^\infty \to \mathbb{R}$, as $u_i\left(\bigcup_{t=0}^{\infty}\{x_{i,j}(t), x_{j,i}(t)\}\right) \triangleq \lim_{t\to\infty} u_{i,t}\left(x_{i,j}(t), x_{j,i}(t)\right)$. When the parameters in the parentheses are clear from the context, we may omit them.

This is the utility we consider in the paper. The utility might be defined otherwise, like a discounted sum, though since we have an exponential convergence, it is possible to simplify it to looking at the limit, assuming that the discounting is not extremely quick. It can be proven that our definition is also equivalent to the other models from Osborne and Rubinstein [21, Chapter 8.3], which are limit of arithmetic means and overtaking. We omit this for lack of space.

2.3 Background

In order to analyze utility in the long run, we use the following convergence theorems from [22], representing what takes place once the actions have stabilized. For two *fixed* agents, they prove:

Theorem 1 *If the reciprocation coefficients are not both 1, which means $r_1 r_2 < 1$, then we have, for $i \in N$: $\lim_{t\to\infty} x_{i,j}(t) = \frac{(1-r_i)k_i + r_i(1-r_j)k_j}{1-r_i r_j}$.*

For two agents, in the *floating* case, they show:

Theorem 2 *If the reciprocation coefficients are neither both 0 and nor both 1, which means $0 < r_1 + r_2 < 2$, then, as $t \to \infty$, $x(t)$ and $y(t)$ converge to a common limit, which is*

$$\frac{1}{2}\left(k_1 + k_2 + (k_2 - k_1)\frac{r_1 - r_2}{r_1 + r_2}\right) = \frac{r_2}{r_1 + r_2}k_1 + \frac{r_1}{r_1 + r_2}k_2.$$

For a *fixed* and a *floating* agent, the following holds:

Theorem 3 *If agent i employs* fixed *reciprocation and the other agent j employs the* floating *one, assume that $r_i < 1$ and $r_j > 0$. Then, both limits exist and are equal to k_i. The convergence is geometrically fast.*

The following holds for two agents with any attitudes:

Proposition 1 *If $k_1 \le k_2$ and both action sequences converge, then $\lim_{t\to\infty} x_{i,j}(t) \le \lim_{t\to\infty} x_{j,i}(t)$.*

3 Utility Maximization

As a first step to analyzing strategic choices, consider how an agent can maximize her utility by choosing either her reciprocation coefficient or reciprocation attitude, before the interaction begins. This can be expected from a rational agent, who reciprocates, but chooses her reciprocation habits. In the case of Example 1, this models a country setting a smart foreign policy with respect to arming. Since in reality the behavioral parameters of others are unknown, choosing an optimal behavior will probably be harder, through trial and error, and the theory predicts the trend of these choices. Some (parts of) proofs are omitted for lack of space.

First, suppose that the only available option of agent i to modify the reciprocation process is by setting its reciprocation coefficient r_i. We therefore analyze how i's utility depends on r_i. In the results of this section, the asymmetry of the agents stems from $k_2 \geq k_1$.

For the *fixed* reciprocation attitude, we prove:

Proposition 2 *In the* fixed *reciprocation attitude, the following holds: If* $r_2 < 1$ *and agent* 1 *wants to maximize his utility by choosing his reciprocation coefficient* r_1*, then he should set* r_1 *to be*
$$\begin{cases} 1 & \text{if } r_2 > \beta_1, \\ \text{anything} & r_2 = \beta_1, \\ 0 & r_2 < \beta_1. \end{cases}$$
If $r_1 < 1$ *and agent* 2 *wants to maximize his utility by choosing his reciprocation coefficient* r_2*, then he should set* r_2 *to be*
$$\begin{cases} 0 & \text{if } r_1 > \beta_2, \\ \text{anything} & r_1 = \beta_2, \\ 1 & r_1 < \beta_2. \end{cases}$$
These choices are the only utility maximizing ones.

The idea of the proof is to express the utility of an agent and differentiate it by her reciprocation coefficient, to find candidates for the extrema.

Proof. Let us prove for agent 1 choosing r_1. We first express 1's utility and then maximize it. Since $r_2 < 1$, we have $r_1 r_2 < 1$, and from Theorem 1,

$$\lim_{t\to\infty} x(t) = \frac{(1-r_1)k_1 + r_1(1-r_2)k_2}{1-r_1r_2},$$
$$\lim_{t\to\infty} y(t) = \frac{(1-r_2)k_2 + r_2(1-r_1)k_1}{1-r_1r_2} \Rightarrow u_1 =$$
$$\frac{(1-r_2)k_2 + r_2(1-r_1)k_1}{1-r_1r_2} - \beta_1 \frac{(1-r_1)k_1 + r_1(1-r_2)k_2}{1-r_1r_2}.$$

To find a maximum point of this utility as a function of r_1, we differentiate:

$$\frac{\partial(u_1)}{\partial(r_1)} = \ldots = \frac{(r_2-\beta_1)(1-r_2)}{(1-r_1r_2)^2}(k_2-k_1).$$

Therefore, if $r_2 = \beta_1$, then the derivative is zero, and the utility is constant. Otherwise, the maximum is attained at an endpoint: at the right endpoint, if the $r_2 > \beta_1$, and at the left endpoint if $r_2 < \beta_1$.

The case of agent 2 choosing r_2 is proven by analogy. □

For the *floating* reciprocation attitude, we prove:

Proposition 3 *In the* floating *reciprocation attitude, the following holds: If* $r_2 < 1$ *and agent* 1 *wants to maximize his utility by choosing his reciprocation coefficient* r_1*, then he should set* r_1 *to be*
$$\begin{cases} 1 & \text{if } r_2 > 0 \text{ and } \beta_1 < 1, \\ 0 & \text{if } r_2 > 0 \text{ and } \beta_1 > 1, \\ \text{anything} & \text{if } r_2 > 0 \text{ and } \beta_1 = 1, \\ 0 & \text{if } r_2 = 0 \text{ and } \beta_1 > 0, \\ \text{anything positive} & \text{if } r_2 = 0 \text{ and } \beta_1 < 0, \\ \text{anything} & \text{if } r_2 = 0 \text{ and } \beta_1 = 0. \end{cases}$$
If $r_1 < 1$ *and agent* 2 *wants to maximize his utility by choosing his reciprocation coefficient* r_2*, then he should set* r_2 *to be*
$$\begin{cases} 0 & \text{if } r_1 > 0 \text{ and } \beta_2 < 1, \\ 1 & \text{if } r_1 > 0 \text{ and } \beta_2 > 1, \\ \text{anything} & \text{if } r_1 > 0 \text{ and } \beta_2 = 1, \\ \text{anything positive} & \text{if } r_1 = 0 \text{ and } \beta_2 > 0, \\ 0 & \text{if } r_1 = 0 \text{ and } \beta_2 < 0, \\ \text{anything} & \text{if } r_1 = 0 \text{ and } \beta_2 = 0. \end{cases}$$
These choices are the only utility maximizing ones.

The idea of the proof is as in the previous proof.

If the kindness values and reciprocation coefficient are set, and an agent may only choose between *fixed* or *floating* reciprocation, we prove:

Proposition 4 *If* $0 < r_1, r_2 < 1$*, then, if agent* 1 *wants to maximize her utility, and she may only choose whether to employ* fixed *or* floating *reciprocation, then she should choose*

$$\begin{cases} \text{fixed} & \text{if } (2 \text{ is fixed} \wedge \{\beta_1 \geq r_2\}) \\ \vee(\text{agent } 2 \text{ is floating} \wedge \{\beta_1 \geq 1\}), & \\ \text{floating} & \text{if } (2 \text{ is fixed} \wedge \{\beta_1 \leq r_2\}) \\ \vee(\text{agent } 2 \text{ is floating} \wedge \{\beta_1 \leq 1\}). & \end{cases}$$

If agent 2 *wants to maximize his utility by choosing* fixed *or* floating *reciprocation, then he should choose*

$$\begin{cases} \text{floating} & \text{if } (1 \text{ is fixed} \wedge \{\beta_2 \geq r_1\}) \\ \vee(\text{agent } 1 \text{ is floating} \wedge \{\beta_2 \geq 1\}), & \\ \text{fixed} & \text{if } (1 \text{ is fixed} \wedge \{\beta_2 \leq r_1\}) \\ \vee(\text{agent } 1 \text{ is floating} \wedge \{\beta_2 \leq 1\}). & \end{cases}$$

Supposing $k_1 < k_2$*, an attitude choice given in this proposition is the only best one if and only if the relevant inequality on the right-hand side of the conditions holds strictly.*

The idea of the proof is to compare the possibilities, to see when which option is best. For $\beta_1 = \beta_2 = 0$, which is when both agents want only to receive more, all the results from this section are intuitive, since a less kind agent should choose to be very reciprocating, while the other agent should choose to be completely non-reciprocating, thereby remaining kind and pulling the other agent to act more.

In Example 1, if countries 1 and 2 have $r_1 = r_2 = 0.5$, $\beta_1 = 0, \beta_2 = 0.2$ (acting is cheap), then, whatever attitude 2 employs, 1 should employ *floating*, to maximize its utility.

We have prepared the analysis of the game of choosing reciprocation habits. To prepare the ground for analyzing the efficiency of NE, our next step will be finding how the social welfare can be maximized.

4 Maximizing Social Welfare

Maximizing the social welfare is relevant for analyzing the whole interaction of agents maximizing their own utilities as a game, to see how good equilibria are for the society relatively to the best possible social welfare. Regardless of the game, the manager (say, the boss of a group of interacting workers) wants to maximize the social welfare by influencing agents' behavior through propaganda or an incentive mechanism.

We now define the social welfare.

Definition 4 *The social welfare at time t* ($\mathrm{SW}_t \colon \mathbb{R}^2 \to \mathbb{R}$) *is defined as the sum of utilities at time t, i.e.,*

$$\mathrm{SW}_t \triangleq u_{1,t} + u_{2,t} = (1-\beta_1)x(t) + (1-\beta_2)y(t). \tag{1}$$

For the whole process, we define the (asymptotic) *social welfare*, $\mathrm{SW} \colon \left(\mathbb{R}^2\right)^\infty \to \mathbb{R}$, as $\mathrm{SW} \triangleq \lim_{t\to\infty} \mathrm{SW}_t$.

In Example 1, changing the behavioral parameters to increase the social welfare models the United Nations trying to spread good practices among countries.

We first suppose that the only available option to influence the interaction network is through choosing the reciprocation coefficients of the agents, and ask what is the most efficient setup of the r_1, r_2 parameters. To this end, we now analyze how the asymptotic social welfare depends on these parameters. Recall that $k_2 \geq k_1$. For given reciprocation attitudes (not necessarily the same attitudes for both agents), we prove

Proposition 5 *We can maximize the social welfare by setting r_1 and r_2 to*

$$\begin{cases} r_1 = 1, r_2 = 0 & \text{if } \max\{\beta_1, \beta_2\} \leq 1, \\ r_1 = 0, r_2 = 1 & \text{if } \min\{\beta_1, \beta_2\} \geq 1, \\ r_1 = r_2 = 0 & \text{if } \beta_1 \geq 1, \beta_2 \leq 1, \\ r_1 = 1, r_2 = 0 & \text{if } \beta_1 \leq 1, \beta_2 \geq 1, \beta_1 + \beta_2 \leq 2, \\ r_1 = 0, r_2 = 1 & \text{if } \beta_1 \leq 1, \beta_2 \geq 1, \beta_1 + \beta_2 \geq 2. \end{cases} \tag{2}$$

The idea of the proof is to consider, what limits should be maximized, to maximize the social welfare.

Proof. If $\max\{\beta_1, \beta_2\} \leq 1$, then if we maximize both $\lim_{t\to\infty} x(t)$ and $\lim_{t\to\infty} y(t)$, we maximize the social welfare. For $r_1 = 1, r_2 = 0$, we obtain[5] $\lim_{t\to\infty} x(t) = k_2$ and $\lim_{t\to\infty} y(t) = k_2$, which are the maximum possible. Thus, $r_1 = 1, r_2 = 0$ maximizes the social welfare.

We skip the easy cases, concentrating on the hard one.

If $\beta_1 \leq 1, \beta_2 \geq 1$, we first express the social welfare in a handier form, and subsequently show how we can maximize it. Denote $\delta \triangleq 1 - \beta_1 \Rightarrow \delta \geq 0$ and $\epsilon \triangleq 2 - \beta_1 - \beta_2$. Then, we have $1 - \beta_2 = -(\delta - \epsilon)$ and $\mathrm{SW} = (1-\beta_1)\lim_{t\to\infty} x(t) + (1-\beta_2)\lim_{t\to\infty} y(t) = \delta \lim_{t\to\infty} x(t) - (\delta - \epsilon)\lim_{t\to\infty} y(t) = \delta(\lim_{t\to\infty} x(t) - \lim_{t\to\infty} y(t)) + \epsilon \lim_{t\to\infty} y(t)$.

Now, if $\beta_1 + \beta_2 \leq 2$, then $\epsilon \geq 0$ and thus, if we maximize $\lim_{t\to\infty} x(t) - \lim_{t\to\infty} y(t)$ and $\lim_{t\to\infty} y(t)$, we maximize the social welfare. For $r_1 = 1, r_2 = 0$, we obtain[5] $\lim_{t\to\infty} x(t) = \lim_{t\to\infty} y(t) = k_2$, thus maximizing the first (since by Proposition 1, $\lim_{t\to\infty} x_{i,j}(t) \leq \lim_{t\to\infty} x_{j,i}(t)$, the first is non-positive) and the second. Thus, $r_1 = 1, r_2 = 0$ maximizes the social welfare.

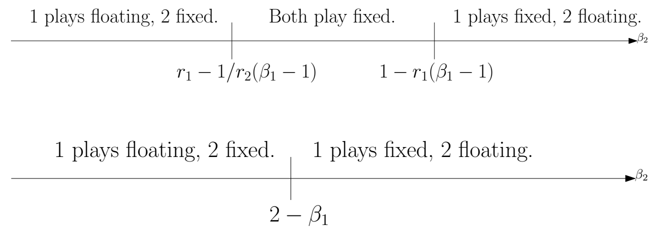

Figure 1: The upper figure is for $\beta_1 - 1 \geq 0$, and the lower figure is for $\beta_1 - 1 < 0$. The strategy profile written above denotes a profile to maximize the social welfare, based on where the value of β_2 resides.

Now, if $\beta_1 + \beta_2 \geq 2$, then $\epsilon \leq 0$ and thus, if we maximize $\lim_{t\to\infty} x(t) - \lim_{t\to\infty} y(t)$ and minimize $\lim_{t\to\infty} y(t)$, we maximize the social welfare. For $r_1 = 0, r_2 = 1$, we obtain[5] $\lim_{t\to\infty} x(t) = \lim_{t\to\infty} y(t) = k_1$, thus maximizing the first and minimizing the second. Thus, $r_1 = 0, r_2 = 1$ maximizes the social welfare. □

Note that this proposition holds also if we may influence both r_1, r_2 and the attitudes of the agents, since the proof maximizes and minimizes expressions for any possible attitudes.

Suppose now that the reciprocation coefficients are set, and the manager only chooses whether the agents employ *fixed* or *floating* reciprocation.

Proposition 6 *If $0 < r_1, r_2 < 1$, then the social welfare is maximal by reciprocating as follows:*

$$\begin{cases} \textit{1 floating, 2 fixed.} & \textit{if } \beta_2 \leq 1 - \max\left\{\frac{1}{r_2}(\beta_1 - 1), \beta_1 - 1\right\}, \\ \textit{1 fixed, 2 fixed.} & \textit{if } 1 - \frac{1}{r_2}(\beta_1 - 1) \leq \beta_2 \leq 1 - r_1(\beta_1 - 1), \\ \textit{1 fixed, 2 floating.} & \textit{if } \beta_2 \geq 1 - \min\{r_1(\beta_1 - 1), \beta_1 - 1\}. \end{cases}$$

The statement of the proposition can be expressed geometrically. We can maximize the social welfare depending on the real interval where β_2 is: Figure 1 shows a profile to maximize the social welfare, based on the segment where the value of β_2 belongs.

The omitted proof compares the various options.

For $\beta_1 = \beta_2 = 0$, this result (agent 1 plays *floating*, 2 *fixed*) is intuitive, since the less kind agent aligns to the kinder one. Also the previous results of this section show that for $\beta_1 = \beta_2 = 0$, the less kind agent should align to the kinder one, to maximize the social welfare. By now, the preparation for analyzing the whole interaction as a game are completed, so we proceed to define and to analyze the game.

5 Reciprocation Attitude Game

We have considered an agent choosing her reciprocation coefficient or her *fixed* or *floating* reciprocation attitude, each choice yielding certain (asymptotic) utility to the agent. This situation is naturally modeled as a game where the strategies of each agent are the above choices and the utility is the asymptotic utility of the interaction. Recall that the utility of agent i is $\lim_{t\to\infty}\{x_{j,i}(t) - \beta_i x_{i,j}(t)\}$. This is a one-shot game, the attitude being chosen once, before the interaction commences. Analyzing this game allows predicting the situation, supplying some advice to an external party (such as the boss who wants to influence her employees) or to the agents

[5] This is evident from the definition of *fixed* or *floating* reciprocation, without a convergence theorem.

themselves. As explained after Example 4, human agents usually neither completely mimic the others' behavior, nor do they completely ignore it, which means $0 < r_1, r_2 < 1$. We call this game the *reciprocation attitude game (RAG)*. Theorems 4 and 5 summarize our findings about RAG.

We first characterize the existence of pure NE in this game and subsequently look into their efficiency. Then, we consider how stable these NE are with respect to the best response dynamics. We assume that $k_2 > k_1$ (strictly) in this section; when the kindness is equal, everyone always keeps acting with this equal value.

Theorem 4 *The NE of RAG are characterized as follows:*

$$\begin{array}{lcl} (\text{fixed}, \text{fixed}) \textit{ is an NE} & \iff & \beta_1 \geq r_2 \textit{ and } \beta_2 \leq r_1. \\ (\text{float}, \text{fixed}) \textit{ is an NE} & \iff & \beta_1 \leq r_2 \textit{ and } \beta_2 \leq 1. \\ (\text{fixed}, \text{float}) \textit{ is an NE} & \iff & \beta_1 \geq 1 \textit{ and } \beta_2 \geq r_1. \\ (\text{float}, \text{float}) \textit{ is an NE} & \iff & \beta_1 \leq 1 \textit{ and } \beta_2 \geq 1. \end{array}$$

The proof utilizes Proposition 4 about utility maximization to see when no deviation is profitable.

Proof. Assume that $\beta_1 \geq r_2$ and $\beta_2 \leq r_1$. If the strategy profile is (*fixed*, *fixed*), then, according to Proposition 4, no agent will have an incentive to unilaterally deviate, meaning this strategy profile is indeed an NE.

Assume now that (*fixed*, *fixed*) is an NE. We prove that $\beta_1 \geq r_2$ and $\beta_2 \leq r_1$ by contradiction. If $\beta_1 < r_2$, then Proposition 4 would imply that agent 1 would like to deviate, contradictory to the profile being an NE. If $\beta_2 > r_1$, Proposition 4 would imply that 2 would like to deviate, contradictory to the NE.

The remaining 3 cases are proven by analogy. □

Remark 1 (Existence of NE) *If no characterizing condition holds, then no NE exists. For example, no characterizing condition holds when $\beta_1 = 0.8, \beta_2 = 0.9, r_1 = 0.5, r_2 = 0.2$, so no pure NE exists in this case. Since the game is finite, a mixed NE always exists by the classical result by Nash [19].*

We now illustrate the theorem for certain parameter values.

Example 5 *Let $\beta_1 = 0.3, \beta_2 = 0.6$. Theorem 4 states that*

$$\begin{array}{lcl} (\text{fixed}, \text{fixed}) \textit{ is an NE} & \iff & 0.3 \geq r_2 \textit{ and } 0.6 \leq r_1. \\ (\text{float}, \text{fixed}) \textit{ is an NE} & \iff & 0.3 \leq r_2. \end{array}$$

No other Nash equilibria exist.

5.1 PoA and PoS

The manager or the government may want to know how far the social welfare in an equilibrium is from the maximum possible social welfare. To this end, we consider the famous measures of the efficiency of an equilibrium, namely price of anarchy [16] (PoA) and price of stability [1] (PoS). PoA is the smallest ratio of a social welfare in an NE to the optimum social welfare, and PoS is the largest such ratio.

Theorem 4 provides all the NE, for each set of parameters. Using Proposition 6, we know for each set of parameters what the maximum social welfare is. Calculating the social welfare at each of the Nash equilibria and finding its ratio to the optimum social welfare enables us to find the price of anarchy and stability in the following theorem.

Conditions:	PoA = PoS :
$1 + r_2 - r_2\beta_2 > \beta_1 > r_2$ $\wedge \{\beta_2 < r_1\}$	$\frac{\sum_{i=1}^{2} (1-\beta_i) \frac{(1-r_i)k_i + r_i(1-r_j)k_j}{1-r_i r_j}}{(2-\beta_1-\beta_2)k_2}$
$\{\beta_1 > 1 + r_2 - r_2\beta_2\} \wedge \{\beta_2 < r_1\}$ $\wedge \{1 + 1/r_1 - \beta_2/r_1 > \beta_1\}$	1
$\beta_1 > 1 + 1/r_1 - \beta_2/r_1$ $\wedge \{\beta_2 < r_1\}$	$\frac{\sum_{i=1}^{2} (1-\beta_i) \frac{(1-r_i)k_i + r_i(1-r_j)k_j}{1-r_i r_j}}{(2-\beta_1-\beta_2)k_1}$
$\{\beta_1 < r_2\} \wedge \{\beta_2 < 1\}$	1
$\{\beta_1 > 1\} \wedge \{1 + r_1 - \beta_1 r_1 > \beta_2\}$ $\wedge \{\beta_2 > \max\{1 + 1/r_2 - \beta_1, r_1\}\}$	$\frac{(2-\beta_1-\beta_2)k_1}{\sum_{i=1}^{2} (1-\beta_i) \frac{(1-r_i)k_i + r_i(1-r_j)k_j}{1-r_i r_j}}$
$\{\beta_1 > 1\} \wedge$ $\beta_2 > \max\{1 + r_1 - \beta_1 r_1, r_1\}$	1
$\{\beta_1 < 1\} \wedge \{2 - \beta_1 > \beta_2 > 1\}$	$\frac{r_2}{r_1+r_2} \frac{k_1}{k_2} + \frac{r_1}{r_1+r_2}$
$\{\beta_1 < 1\} \wedge \{\beta_2 > 2 - \beta_1\}$	$\frac{r_2}{r_1+r_2} + \frac{r_1}{r_1+r_2} \frac{k_2}{k_1}$

Table 1: The efficiency of NE in reciprocation attitude game.

Theorem 5 *The efficiency of the equilibria is given in Table 1. In the case of equality in the conditions, the highest entry from our conditions that border the equal value is the price of stability, and the lowest entry is the price or anarchy.*

In particular, if $\beta_1 < r_2, \beta_2 < 1$, then PoA = PoS = 1. We now illustrate the efficiency ranges on Example 5.

Example 5 (Continued) *Recall that $\beta_1 = 0.3, \beta_2 = 0.6$. For these values, Theorem 5 implies the following.*

Conditions:	*Price of anarchy and stability:*
$\{0.3 > r_2\}$ *and* $\{0.6 < r_1\}$	$\frac{\sum_{i=1,2; j \neq 1} (1-\beta_i) \frac{(1-r_i)k_i + r_i(1-r_j)k_j}{1-r_i r_j}}{1.1 k_2}$
$\{0.3 < r_2\}$	1

Consider Example 2. If agents 1 and 2 have $r_1 = r_2 = 0.5$, $\beta_1 = 0, \beta_2 = 0.2$ (acting is cheap), then, as just mentioned, PoA = PoS = 1 and the only NE is (*float*, *fixed*). This is intuitive, since agent 1 will align to the kinder 2, thereby each agent maximizes the total action and, since acting is cheap, also her own utility and the social welfare.

This completes the analysis of the agents setting their own reciprocation attitudes. The next section considers agents who set both their own reciprocation attitudes and coefficients.

6 Reciprocation Attitude and Coefficient Game

In the previous section we looked at the game of choosing a reciprocation attitude. It is also natural to consider what happens when the other habit, namely, the reciprocation coefficient, is chosen as well. Analyzing this game allows predicting the situation of more choice than the situation analyzed in RAG; for instance, the participants have more willpower or knowledge than in RAG. As before, this is a one-shot game, the attitude and reciprocation coefficient being chosen once, before the interaction commences. As we did for RAG, since people usually neither completely mimic the others' behavior, nor do they completely ignore it, we assume $0 < r_1, r_2 < 1$. We call this game the *reciprocation attitude and coefficient game (RACG)*. This game is analyzed in Theorems 6 and 7.

We first characterize the existence of pure NE in this game and then look into their efficiency, by finding the price of anarchy and stability. This section assumes that $k_2 > k_1$ (strictly).

Theorem 6 *The only Nash equilibria of RACG are characterized as follows:*

$$\begin{array}{ccc}
\textit{An equilibrium profile:} & & \textit{Condition:}\\
(\text{fixed, fixed, } r_1 = \beta_2, r_2 = \beta_1) & \iff & 0 < \beta_1, \beta_2 < 1.\\
(\text{float, fixed, } 0 < r_1, r_2 < 1, \beta_1 \leq r_2) & \iff & \beta_1 < 1 \wedge \beta_2 \leq 1.\\
(\text{fixed, float, } 0 < r_1, r_2 < 1, r_1 \leq \beta_2) & \iff & \beta_1 \geq 1 \wedge \beta_2 > 0.\\
(\text{float, float, } 0 < r_1, r_2 < 1) & \iff & \beta_1 = \beta_2 = 1.
\end{array}$$

The proof is based on Theorem 4, which narrows down the set of possible Nash equilibria, on Proposition 2 and Proposition 3 about utility maximization, and on convergence results from [22] (See Section 2.3.)

Proof. We go over all the NE for RAG from Theorem 4 and look at all the possible choices of r_1 and r_2 to have an equilibrium in the new game. No other equilibria exist, since if no condition of Theorem 4 is satisfied, then even deviating by changing only the attitude is possible.

We begin with (*fixed*, *fixed*), an NE in RAG if and only if $\beta_1 \geq r_2$ and $\beta_2 \leq r_1$. Given these reciprocation attitudes, Proposition 2 implies that to prevent the only best choice of r_1 being 0 or 1, we must have $(r_2 - \beta_1) = 0$, and to avoid the situation where the only best choice of r_2 is 0 or 1, we must have $(\beta_2 - r_1) = 0$. This implies the necessity of the conditions for an NE with *fixed* attitudes. Theorem 4 and Proposition 2 imply that these conditions are also sufficient to prevent deviations of only the attitude or only the reciprocation coefficient. If agent j simultaneously deviates to another attitude and r_j, then Theorem 3 implies that any $r_j > 0$ yields the same utility, and therefore, this deviation may be considered to consist of attitude only, which is known to be not profitable. This proves the sufficiency.

Consider the profile (*float*, *fixed*) now, an NE in RAG if and only if $\beta_1 \leq r_2$ and $\beta_2 \leq 1$. Since $r_2 < 1$ implies that $\beta_1 \leq r_2 < 1$, we have the necessity of the conditions for a NE with *floating* and *fixed* attitudes. Theorem 4 implies that deviating in attitude only is not profitable. By Theorem 3, any $r_1, r_2 \in (0, 1)$ suffice for a best response, and so deviating in reciprocation coefficient only is not profitable as well. Consider a deviation of an agent to another attitude and reciprocation coefficient simultaneously. Unless this includes r_2 becoming less than β_1, we still know from what we have just proven that for this new profile, a deviation by the attitude only would not benefit agent 2, and since changing r_2 has not been profitable, the whole deviation is not profitable. The only remaining option is agent 2 becoming *floating* and changing r_2 to be less than β_1. This would yield agent 2 the utility of $(1 - \beta_2)(\frac{r_2}{r_1+r_2}k_1 + \frac{r_1}{r_1+r_2}k_2)$, by Theorem 2, while he previously had, by Theorem 3, $(1 - \beta_2)k_2$. Since $1 - \beta_2 \geq 0$ and $k_2 > k_1$, the previous profit is not smaller than the new one.

The two remaining cases are similar. □

Remark 2 (Existence of NE) *When no characterizing condition holds, no NE exists. For instance, if $\beta_1 < 1 < \beta_2$, no characterizing condition holds, and therefore, no (pure) NE exists.*

6.1 PoA and PoS

We now look at the efficiency of these equilibria, proving

Theorem 7 *The efficiency of the NE is given in Table 2.*

We find the possible NE from Theorem 6, and compare their social welfare with the optimal social welfare, found based on the proof of Proposition 5. We only use the ideas of what one should minimize or maximize to maximize the social welfare from the proof of Proposition 5, since the proposition sets reciprocation coefficients to 0 and 1, so we cannot use it directly. To calculate the social welfare, we use the definition of utility and the limit values from Theorems 1, 2, and 3.

Proof. If $0 < \beta_1, \beta_2 < 1$, Theorem 6 implies that there exist exactly two Nash equilibria, namely (*fixed*, *fixed*, $r_1 = \beta_2, r_2 = \beta_1$) and (*float*, *fixed*, $0 < r_1, r_2 < 1, \beta_1 \leq r_2$). For the optimal social welfare, we need to maximize both $\lim_{t\to\infty} x(t)$ and $\lim_{t\to\infty} y(t)$, as does, for instance, the second NE above, yielding the social welfare of $(2-\beta_1-\beta_2)k_2$. Taking the ratios of the social welfare values gives row one in the table from the statement of the theorem.

The remaining cases are proven using the same idea. □

For an RAG, Theorem 5 implies that small enough β_1, β_2 guarantee that all the NE are optimal. In RACG, however, when $0 < \beta_1, \beta_2 < 1$, the proof of Theorem 7 shows that along with a socially optimal NE, the social welfare of the NE (*fixed*, *fixed*, $r_1 = \beta_2, r_2 = \beta_1$) relative to the optimum is $\frac{\sum_{i=1,2;j\neq i}(1-\beta_i)\frac{(1-\beta_j)k_i+\beta_j(1-\beta_i)k_j}{1-\beta_j\beta_i}}{(2-\beta_1-\beta_2)k_2}$. When the efforts of acting approach zero for both agents, this expression approaches

$$\lim_{\beta_1\to 0,\beta_2\to 0} \frac{\sum_{i=1,2;j\neq i}(1-\beta_i)\frac{(1-\beta_j)k_i+\beta_j(1-\beta_i)k_j}{1-\beta_j\beta_i}}{(2-\beta_1-\beta_2)k_2}$$
$$= \frac{\sum_{i=1,2;j\neq i} k_i}{2k_2} = \frac{k_1+k_2}{2k_2} = \frac{1}{2}(\frac{k_1}{k_2}+1).$$

That is, allowing more freedom (setting own reciprocation attitude and coefficient), we may lose up to half of the efficiency, if k_1/k_2 is small. However, Theorem 7 leaves a sparkle of hope: if at least one agent acts completely effortlessly or even enjoys it, meaning that $\beta_i \leq 0$, then all the NE are socially optimal.

We now turn to the case of n agents, being done with 2.

7 Arbitrarily Many Agents

The original model of [22] is defined for any number $n \geq 2$ of reciprocating agents, where every agent has both r_i and r'_i, the second reciprocation coefficient being the fraction of the action, determined by reacting to the average of all the other agents' actions. They prove convergence, but find the limit only when all the agents are *floating*, and that is the technical obstacle to generalize this paper to n agents. We can, however, assume n *floating* agents and analyze the game of choosing only the reciprocation coefficient, called *reciprocation coefficient game*, by finding its equilibria and their efficiency.

In this case, we discover again that when acting is easy ($\beta_i = 0$), then the kinder agents should pull the less kind ones to act more while not reacting much to the actions they receive by acting less. The results also imply that if all the $1 - \beta_i$s have the same sign, then PoA $= 1$.

Conditions:	Price of anarchy:	Price of stability:
$0 < \beta_1, \beta_2 < 1$	$\frac{\sum_{i=1,2;j \neq i} (1-\beta_i) \frac{(1-\beta_j)k_i + \beta_j(1-\beta_i)k_j}{1-\beta_j \beta_i}}{(2-\beta_1-\beta_2)k_2}$	1
$\beta_1 < 1$ and $\beta_2 \leq 1$ but not $0 < \beta_1, \beta_2 < 1$	1	1
$\beta_1 \geq 1$ and $0 < \beta_2 \leq 1$ but not $\beta_1 = \beta_2 = 1$	$\frac{(1-\beta_1-\beta_2)k_1}{(1-\beta_1)k_1+(1-\beta_2)k_2}$	$\frac{(1-\beta_1-\beta_2)k_1}{(1-\beta_1)k_1+(1-\beta_2)k_2}$
$\beta_1 \geq 1$ and $\beta_2 > 1$	1	1
$\beta_1 = \beta_2 = 1$	1	1

Table 2: The efficiency of NE for a reciprocation attitude and coefficient game.

In addition to the just described game, there exist many other variations, even for two agents. For instance, for two agents we are able to analyze the game of choosing the reciprocation coefficient for the not *floating* case too. Another variation would be choosing the reciprocation coefficient in a closed segment $[a, b]$, for any $0 < a < b < 1$. This would limit the domain, but the compactness of the domain may facilitate existence of NE. On the other hand, allowing the extreme points $r_i = 0$ or 1 with a proper handling of the cases of no convergence is also an alternative. We can never cover every possible model, but we believe our model sheds light on the general phenomena.

8 Converging to NE

To analyze the stability of Nash equilibria, we recall the famous best response dynamics [21, Section 2.2], where each agent best responds to the current profile of the others. A reasonable question is when and whether this process converges to a NE. For reciprocation attitude games, we prove that given a NE and any profile, we can let each agent simultaneously choose her reciprocation attitude to maximize her utility, such that it ends up in this NE. The same can be proven for reciprocation coefficient games, described in Section 7. For reciprocation attitude and coefficient game, however, the non-compactness of the domain does not allow a best response to always exist. Therefore, the best response process may be undefined. Details are omitted for lack of space.

9 Conclusions and Future Work

We aim to predict and advise on strategic behavior in reciprocation, in both human-human and human-machine interactions. A reciprocal action is modeled as a balance between the inner self and a reaction to others' actions. We define an agent's utility asymptotically. We then consider choosing the reciprocation attitude or coefficient to maximize her own utility. We finally model the strategic behavior of the reciprocating agents in several games, characterize the NE and their efficiency. We also show that NE may always be achieved by a natural process, the best response dynamics [21, Section 2.2], besides in a RACG. This gives hope for achieving a situation that is stable to unilateral deviations without any regulation.

Our main advice is that both for maximizing own utility and for maximizing the social welfare, if contributing is cheaper than receiving, then, both in choosing the reciprocation attitude and coefficient, the kinder agent should be most stable (be *fixed* or have the reciprocation coefficient $r_i = 0$), and the opposite should be done if contributing is costlier than receiving. When contributing is much cheaper than receiving (β_is are smaller than all the other parameters), then, for the reciprocation attitude game and for the reciprocation coefficient game, the price of anarchy is 1, so rationally reciprocating agents will play socially optimally. In such equilibria, the kinder agents are stable and the less kind agents follow the kinder ones. For the reciprocation attitude and coefficient game, the price of stability is 1, but the price of anarchy is positive, meaning that rationally reciprocating agents may play socially optimally, but may also play suboptimally, so that coordination would be useful.

Comparing Theorem 5 for choosing only the reciprocation attitudes to Theorem 7 for choosing the coefficients as well, we observe that more freedom of choice allows for a socially suboptimal equilibrium, achieving as little as about half of the optimal social welfare, if the kindness values are very different. This pitfall emphasizes the importance of cooperation, if more freedom and power lies at our disposal. Like Churchill said[6]: "Where there is great power there is great responsibility".

The analysis also relates to some real-life phenomena. Our results regarding maximizing utility and social welfare show why in life, if acting is not too hard, then following the example of the kindest makes the individuals and the society thrive, which has already been observed [2]. Since being polite usually consists of words and simple gestures, and is therefore quite easy for many people, this explains why people choose this strategy with experience, becoming more polite, as is indeed observed [12]. In diplomacy (Example 1), these results predict that diplomats will be polite to each other, since this does not take much effort. Being polite benefits the individual and the society by making people feel better easily.

Many interesting directions for further research exist: *a*) Modeling changes in the reciprocity coefficients, attitudes, or βs during the interaction and not only before it starts. *b*) Modeling probabilistic reaction. *c*) Looking how the manager can really influence the behavior of the agents. *d*) Real agents often join and leave the interaction dynamically. For example, people get born and immigrate to a country, some die and emigrate. Therefore, dynamic interaction is very interesting. *e*) We used others' research, based on real data, as a basis for the model. Therefore, verifying the model on relevant data, like the arms race actions, would be interesting.

Our analysis provides behavioral advice and predicts reciprocation phenomena. It lays the foundation for further modeling of reciprocation, required to even better anticipate and improve the individual utilities and the social welfare.

ACKNOWLEDGEMENTS

This work has been supported by the project SHINE, the flagship project of DIRECT (Delft Institute for Research on ICT at Delft University of Technology).

[6] This quote is from the French National Convention, 08/05/1793.

REFERENCES

[1] E. Anshelevich, A. DasGupta, J. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden, 'The price of stability for network design with fair cost allocation', in *Foundations of Computer Science, 2004. Proceedings. 45th Annual IEEE Symposium on*, pp. 295–304, (Oct 2004).

[2] R Axelrod and WD Hamilton, 'The evolution of cooperation', *Science*, **211**(4489), 1390–1396, (1981).

[3] R.M. Axelrod, *The evolution of cooperation*, Basic books, Basic Books, 1984.

[4] Robert Axelrod, 'The emergence of cooperation among egoists', *American Political Science Review*, **75**, 306–318, (jun 1981).

[5] James C. Cox, Daniel Friedman, and Steven Gjerstad, 'A tractable model of reciprocity and fairness', *Games and Economic Behavior*, **59**(1), 17 – 45, (2007).

[6] Ariel Rubinstein Dilip Abreu, 'The structure of nash equilibrium in repeated games with finite automata', *Econometrica*, **56**(6), 1259–1281, (1988).

[7] William J. Dixon, 'Reciprocity in united states-soviet relations: Multiple symmetry or issue linkage?', *American Journal of Political Science*, **30**(2), pp. 421–445, (1986).

[8] Martin Dufwenberg and Georg Kirchsteiger, 'A theory of sequential reciprocity', *Games and Economic Behavior*, **47**(2), 268 – 298, (2004).

[9] Armin Falk and Urs Fischbacher, 'A theory of reciprocity', *Games and Economic Behavior*, **54**(2), 293 – 315, (2006).

[10] Ernst Fehr, Urs Fischbacher, and Simon Gchter, 'Strong reciprocity, human cooperation, and the enforcement of social norms', *Human Nature*, **13**(1), 1–25, (2002).

[11] Ernst Fehr and Simon Gächter, 'Fairness and retaliation: The economics of reciprocity', *Journal of Economic Perspectives*, **14**(3), 159–181, (2000).

[12] Maria Rosa Baroni Giovanna Axia, 'Linguistic politeness at different age levels', *Child Development*, **56**(4), 918–927, (1985).

[13] John Gottman, Catherine Swanson, and James Murray, 'The mathematics of marital conflict: Dynamic mathematical nonlinear modeling of newlywed marital interaction', *Journal of Family Psychology*, **13**, 3–19, (1999).

[14] Werner Güth, Rolf Schmittberger, and Bernd Schwarze, 'An experimental analysis of ultimatum bargaining', *Journal of Economic Behavior & Organization*, **3**(4), 367 – 388, (1982).

[15] D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, 2011.

[16] E. Koutsoupias and C. Papadimitriou, 'Worst-case equilibria', in *16th Annual Symposium on Theoretical Aspects of Computer Science*, pp. 404–413, Trier, Germany, (4–6 March 1999).

[17] W.F.G. Mastenbroek, *Onderhandelen*, Het Spectrum, 1992.

[18] Heather McAllister, *Who You Are Is What You Do: Making Choices About Life After School*, Wilkins Farago Pty Ltd, illustrated edn., February 2013.

[19] John Nash, 'Non-Cooperative Games', *The Annals of Mathematics*, **54**(2), 286–295, (September 1951).

[20] N. Nisan, T. Roughgarden, E. Tardos, and V.V. Vazirani, *Algorithmic Game Theory*, Cambridge University Press, 2007.

[21] Martin J. Osborne and Ariel Rubinstein, *A Course in Game Theory*, volume 1 of *MIT Press Books*, The MIT Press, April 1994.

[22] G. Polevoy, M.M. de Weerdt, and C.M. Jonker, 'The convergence of reciprocation', in *Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '16, Richland, SC, (2016). International Foundation for Autonomous Agents and Multiagent Systems.

[23] Matthew Rabin, 'Incorporating fairness into game theory and economics', *The American Economic Review*, **83**(5), pp. 1281–1302, (December 1993).

[24] Ariel Rubinstein, *Modeling Bounded Rationality*, volume 1, The MIT Press, 1 edn., 1997.

[25] Hayley Spencer, 'Why reciprocation is key to building business relationships', *Business 2 Community*, (2012).

[26] Robert L. Trivers, 'The evolution of reciprocal altruism', *The Quarterly Review of Biology*, **46**, 35–57, (mar 1971).

[27] M. Don Ward, 'Modeling the USA-USSR arms race', *Transactions of The Society for Modeling and Simulation International*, **43**, 196–203, (1984).