Theoretical models of decision-making in the Ultimatum Game: Fairness vs. Reason.

Tatiana V. Guy[1], Miroslav Kárný[1],
Alessandra Lintas[2], and Alessandro E.P. Villa[2]

[1] Department of Adaptive Systems,
Institute of Information Theory and Automation,
The Czech Academy of Sciences, Prague, Czech Republic
[guy, school]@utia.cas.cz
[2] Neuroheuristic Research Group & LABEX, HEC Lausanne,
University of Lausanne, Quartier UNIl-Dorigny,
CH-1015 Lausanne, Switzerland
[avilla,alintas]@unil.ch

Abstract. According to Game Theory a human subject playing the Ultimatum Game should choose more for oneself and offer the least amount possible for co-players (*assumption of selfish rationality*) [1]. However, economy, sociology and neurology communities repeatedly claim *non-rationality* of the human behaviour [2], following the observation that responders reject offers they find too low and proposers often offer more than the smallest amount, thus suggesting that humans' behaviour is significantly influenced by social norms. We also assume human rationality, but our model describes a human-responder via decision process with a reward function respecting fairness as much as the economic profit. This model is positively tested against a set of original experimental data, thus providing an insight into human's motivation as a *social* being.

1 Introduction

Decision-making (DM) is considered the most essential phase in a human volitional act and according to traditional economic models [3] humans could be replaced by "rational agents" described as "cold gain maximizer" [4]. Predictions implied by this are well seen on the considered Ultimatum Game (UG). In the UG [5] two players have to split a sum (say 10CHF), with one acting as the proposer and the other as the responder. If the responder accepts the offer, the money is split accordingly. If the responder refuses, both players gain nothing. The rational DM strategy, suggested by the Game Theory, predicts that the two players' behaviours should converge towards the Nash equilibrium. The best decisions include accepting even the smallest possible offer. In reality, proposers tend to offer rather fair offers and responders' tend to reject offers that are judged as unfair (e.g., less than 20 percent of the shared amount).

An intuitively plausible interpretation of this phenomenon is that responders would rather give up some profit than be treated unfairly. This behaviour provides an insight into human's motivation as a *social* being [6].

The aim of this paper is to present a model that considers fairness aspects as the cause of the deviations from the predicted game-theoretical behaviour in UG responder's behaviour. The model shows that apparent irrational behaviour is indeed rational if reward functions include social factors of decision-making such as human attitude to fairness. The proposed model is tested against a set of real data and extremely well predicts responders' decisions.

2 Theoretical and Experimental Methods

Ultimatum game as a decision-making problem In real-life we can assume that humans are driven to maximize their gain iteratively in a sequence of transactions. This kind of situation occurs with an iterative implementation of the UG [7], which is originally a one-shot bargaining game with no communication and no negotiation. The first player (*Proposer*) offers how to split a limited resource (an amount of money q). If the second player (*Responder*) accepts the deal, the resource is distributed according to the proposal (that is s for *Responder* and $(q-s)$ for *Proposer*), otherwise both players get nothing.

Let us consider a DM problem, where the decision maker is *Responder*, while the *Proposer* is a stochastic process. In these settings the *observed state* $s \in \mathcal{S}$ is an offer, *decision* $d \in \mathcal{D}$, where $\mathcal{D}$ ={rejection, acceptance}. The aim is to find a *Responder*'s strategy maximising *Responder*'s economic profit and attitude to fairness. After $i \in \mathbb{N}$ rounds *Responder*'s profit is $z_R(i) = \sum_{k=1}^{i} s_k(d_k-1)$ and *Proposer*'s profit is $z_P(i) = \sum_{k=1}^{i}(q-s_k)(d_k-1)$, where $s_i \in \mathcal{S}$, $\mathcal{S} = \{1, \dots, q-1\}$, is an offer in the ith round, and $d_i \in \mathcal{D} = \{1, 2\}$ is a decision in the ith round, where 1 denotes rejection, 2 stands for acceptance.

Model of Responder Let *Responder*'s reward at the ith round be defined by $r_i(s, d)$ and a weight $w \in [0, 1]$ be associated with *Responder*'s fairness. We considered three alternative types of *Responder*:

R0: ***Rational Responder*** who follows Game Theory. The optimal decision is to accept any non-zero offer. *Responder*'s reward at the ith round equals $r_i(s, d) = z_R(i)$, where $z_R(i)$ is a *pure economic profit*.

R1: ***Mutual Fairness Responder*** who cares about fairness for both players. For this type of *Responder*, the reward at the ith round is modelled by $r_i(s, d) = wz_R(i) - (1-w)\,|z_R(i) - z_P(i)|$, where $w\,z_R(i)$ is *Responder*'s weighted economic profit and the term $(1-w)|z_R(i) - z_P(i)|$ reflects *mutual fairness*. The corresponding optimal decisions were obtained via a simple maximization of the current reward. This is a greedy approximation of the optimal DM.

R2: ***Selfish Fairness Responder*** who cares about fairness only towards himself. *Responder*'s reward at the ith round is $r_i(s, d) = w\,z_R(i) - (1-w)\,z_P(i)$, where $w\,z_R(i)$ is a *weighted economic profit* and $(1-w)\,z_P(i)$ reflects *selfish* *fairness*. In this case the optimal decisions can be found explicitly and are $d_i = 2$ ("accept") if $s_i \geq (1-w)q$ and $d_i = 1$ ("reject"), otherwise.

Learning of Weights Learning of weights is performed in real-time for each *Responder*. The learning algorithm is the same for **R1** and **R2**.
Assuming *Responder*'s rationality, *Responder* decisions are optimal with respect to the reward containing both economic profit and fairness. The degree of balance between these two components (expressed by the weight w in **R1** and **R2**) is specific for any *Responder* and can be learned from the decisions made.

At ith round available data consists of $i-1$ past offers and *Responder*'s decisions (accept or reject). Assumption on optimality of *Responder*'s past decisions implies $i-1$ linear inequalities giving lower $\underline{w}_{i-1}$ and upper $\overline{w}_i$ bounds on weights fitting this assumption. The centre of the interval $[\underline{w}_{i-1}, \overline{w}_{i-1}]$ serves as the current weight estimate. The optimal decision, made for the weight estimate and the offer s_i, serves as a prediction of *Responder*'s decision.

The accuracy of exactly predicted decisions from all made by *Responder*'s then indicates how much the *Responder* acts as the optimiser of the reward considered. This is the specific leave-one-out validation of *Responder* model.

Participants and Behavioural Procedure Twenty neurological healthy, right-handed participants (of either sex, age range $18-45$) volunteered to participate in the study and played with virtual money. All had normal or corrected-to-normal vision and all were naive to the UG. They were informed about the UG test at the beginning of the study and provided consent for their participation in line with the Declaration of Helsinki and upon approval by the ethics committee of the Faculty of Business and Economics of the University of Lausanne. The participants were comfortably seated in a sound- and light-attenuated room. The task was implemented using the E-Prime software (Psychology Software Tools, Inc., Sharpsburg, PA 15215-2821, USA). The participants watched a computer-controlled 19" LCD monitor, with SXGA resolution at a distance of $50-60$ cm. They were instructed about the task and to maintain their gaze on the central fixation cross throughout the experiment. The numerical keypad was used as a response device. The experiment consisted of one block beginning with 20 practice trials to familiarize participants with the task. Each participant played both roles of proposer (90 trials overall) and responder (90 trials overall, Fig. 1) in three alternated blocks of 30 trials each. Participants were told to play the UG trying to maximize their profit as much as possible, irrespective of their role in the game. Each UG trial involved a take-it-or-leave-it integer split of 10CHF. Participants in this study played against a second player that was in fact a computer program (virtual player), even though participants were not told explicitly (task instructions mentioned a generic "second player").

Each "responder" trial started with the pressure of the spacebar of the computer keyboard (event B at time 0, Fig. 1). In this case the proposer, the virtual player, implemented a strategy such that offers occurred randomly with an equal

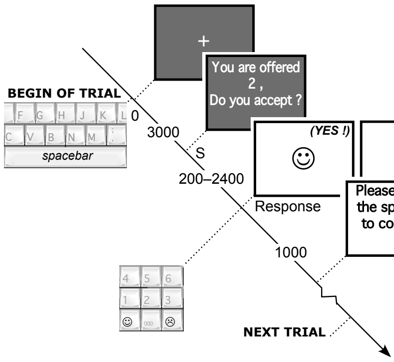

Fig. 1. Illustration of the Ultimatum Game task with the participants acting as responders. Event (S) indicate the stimulus onset. Time intervals are indicated in ms.

frequency of 14.28% each for values in the range $3 - -7$ and with an equal frequency of 7.15% each for values 1, 2, 8, or 9. After an interval of 3000 ms during which participants maintained their gaze on the central fixation cross the message "You are offered s. Do you accept ?", corresponding to event S, appeared on the center of the monitor. The responder's decision (event HR, human player response, Fig. 1) was conveyed by pressing the bottom left key (YES), labeled with a smiled face smiley, of the numerical keypad in case of acceptance and by pressing the bottom right key (NO), labeled with a frowned face smiley, in case of rejection of the offer. An additional 1000 ms interval followed until the message "Please press the spacebar to continue" appeared on the center of the monitor. By pressing the spacebar a new responder trial started. At the end of the block of sessions, the participant was informed about the responder's cumulative profit with a message "Your total gain so far is z CHF".

3 Results

The models were tested against each human participant as follows: at each trial $i \in \{1, \ldots, 90\}$, the weight w_i was dynamically updated using the last offer s_{i-1} and decision d_{i-1}. Then, for the learned weights, a prediction of *Responders*'s decision for the current offer s_i was computed independently for each model **R0**, **R1**, **R2**. The overall rate of accuracy (between 0 and 1) for the three models is described in Table 1.

Table 1. Accuracy rate for three models applied to a sample of 20 human participants playing the Ultimatum Game. Mean, maximal, minimal values and SEM.

UG Responder's model	Mean	min	max	SEM
R0: Rational Responder	0.6640	0.3596	0.9438	0.0310
R1: Mutual Fairness Responder	0.5107	0.1011	0.9213	0.0543
R2: Selfish Fairness Responder	0.8730	0.6067	0.9438	0.0158

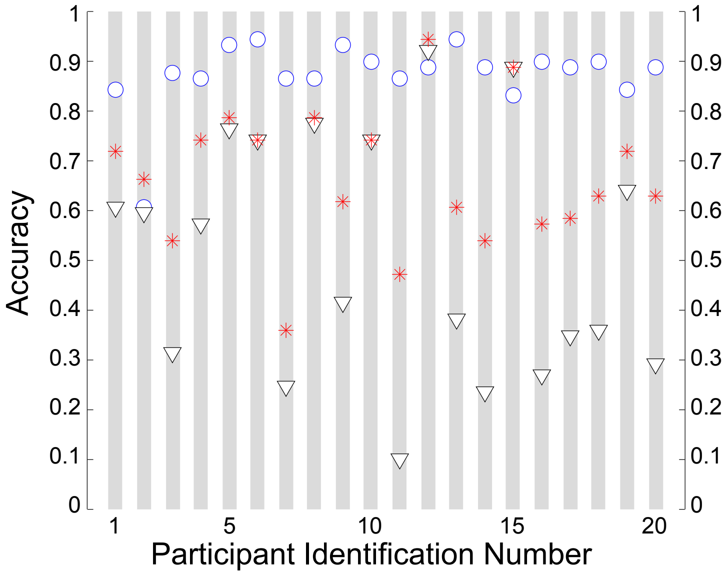

Fig. 2. Accuracy rate of predictions made by **R0** (*), **R1** (∇) and **R2** (**o**) models for each participant of the experimental sample.

The accuracy rate of all models for each participant is illustrated in Fig. 2. We observed that in 17/20 participants model **R2** was significantly the best predictor. In the remaining 3 cases no model was significantly better than the others. Interestingly we observed that in almost half of cases **R0** and **R1** provided similar rate of predictions, although **R0** tended to perform better than **R1**.

4 Discussion

This paper considers rationality of a human-responder in UG. It studies new models able to account responder's behaviour [8, 9]. The key idea is that human rationality is based on a complex reward that includes a social profit as well as the expected economic profit. The balance between economic and social terms is expressed by the responder's attitude to fairness of sharing an amount in UG.

The performed evaluations dynamically estimate human-responders attitude to fairness and predicts the next decision of a human responder using the learned attitude values from the previous trials. This means that there is an incremental learning of the model. The comparison of the actual decisions made by the human responders and the predicted decisions made by the models has shown that the selfish fairness responder (**R2**) was performing much better than the others (prediction accuracy rate of $87.3 \pm 1.6\%$).

The results obtained confirms the hypothesis about rationality of a human-responder in UG with the reward function including selfish player's sense for fairness. In agreement with several previous studies reported in the literature we confirmed that cold gain maximizers (rational responders **R0** according to Game Theory) could not provide a satisfactory level of prediction ($66.4 \pm 3.1\%$). The model **R1**, including fairness equally for *Responder* and for *Proposer*, performed similar to the other models only for 4 participants, thus suggesting that mutual fairness is generally discarded by human responders.

In a multi-player environment concepts like "fairness" and "social sharing" involve the description of an emotional event by the person who experienced it to another person in a socially-shared language [10]. A "fair" share is "irrationally" expected by the participants and they will accept nothing less. Emotions are powerful drives that affect the decision to accept or reject a monetary offer. Following our purely behavioural preliminary results [11] we plan to update the proposed model beyond the fairness framework presented here in order to respect emotional state of the human.

Acknowledgments

The work was supported by the project CAČR 13-13502S of the Czech Science Foundation and by the Swiss National Science Foundation grant CR13I1-138032.

References

1. Rubinstein, A.: Perfect equilibrium in a bargaining model. Econometrica **50**(1) (1982) 97–109
2. Werner, G., Schmittberger, R., Schwarze, B.: An experimental analysis of ultimatum bargaining. J Econ Behav Organ **3**(4) (1982) 367–388
3. Von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press (1944)
4. Thaler, R.H.: From Homo economicus to Homo sapiens. Journal of Economic Perspectives **14** (2000) 133–141
5. van't Wout, M., Chang, L.J., Sanfey, A.G.: The influence of emotion regulation on social interactive decision-making. Emotion **10**(6) (Dec 2010) 815–821
6. Binmore, K.G.: Game Theory and the Social Contract: Just playing. Volume 2. Cambridge : MIT Press (1998)
7. Güth, W., Schmittberger, R., Schwarze, B.: An experimental analysis of ultimatum bargaining. J Econ Behav Organ **3**(4) (1982) 367–388
8. Knejflová, Z., Avanesyan, G., Guy, T., Kárný, M.: What lies beneath players' non-rationality in ultimatum game? In Guy, T., Kárný, M., eds.: Proc. of the 3rd Int. Workshop on Scalable Decision Making, ECML/PKDD 2013
9. Guy, T., Kárný, M., Wolpert, D.: Decision Making with Imperfect Decision Makers. Volume 28. Springer, Berlin (2012)
10. Lane, A., Luminet, O., Rimé, B., Gross, J.J., de Timary, P., Mikolajczak, M.: Oxytocin increases the willingness to socially share one's emotions. International Journal of Psychology (2012)
11. Fiori, M., Lintas, A., Mesrobian, S., Villa, A.E.P.: Effect of emotion and personality on deviation from purely rational decision-making. In Guy, T., Kárný, M., Wolpert, D., eds.: Decision Making and Imperfection. Volume 474. Springer, Berlin (2013) Studies in Computation Intelligence.