The Social Ultimatum Game

Yu-Han Chang, Tomer Levinboim, Rajiv Maheswaran
Department of Computer Science, Information Sciences Institute
University of Southern California, Los Angeles, CA 90292
ychang@isi.edu, tomer.levinboim@usc.edu, maheswar@usc.edu

Abstract

The Ultimatum Game is a key exemplar that shows how human play often deviates from "rational" strategies suggested by game-theoretic analysis. One explanation is that humans cannot put aside the assumption of being in a multi-player multi-round environment that they are accustomed to in the real world. We introduce the Social Ultimatum Game, a multi-player multi-round version of the classical Ultimatum Game. We develop mathematical models of human play that include "irrational" concepts such as fairness and adaptation to societal expectations. We also investigate the stability of maintaining a society of "fair" agents under these conditions. This work is a first step towards building a general theory of imperfect, but reasonable, human-like strategic play in repeated multi-agent games.

1 Introduction

The Ultimatum Game has been studied extensively over the last three decades and is a prominent example of how human behavior deviates from game-theoretic predictions that use the "rational actor" model. The classical game involves two players who are given the opportunity to split $10. One player proposes a potential split, and the other can accept, in which case the players receive the amounts in the proposal, or reject, in which case, both players receive nothing. In the subgame perfect Nash equilibrium, the first player offers $1 to the other player and keeps $9, and the second player accepts the $1 offer, because $1 is better than nothing. However, when experiments are conducted with human subjects, this behavior is rarely observed.

One seemingly intuitive explanation that has not received much treatment in the literature is that humans play similar games in their real lives, and may not view the experimenter's game independently of these more familiar situations. When faced with an isolated Ultimatum Game in the lab, humans thus act in the way that is habitual to them given a similar real-life game. One key feature of these contextual experiences is that humans repeatedly interact over time with multiple other players. The strategy space in such repeated, multi-player games in much more complex, and introduces many new equilibrium strategies. The "irrational" behavior observed in isolated Ultimatum Game experiments no longer seems so irrational when viewed through this lens.

This work is the first step towards building a general theory of imperfect, but reasonable, strategic play in repeated, multi-agent games. By building a model that describes human play in a well-studied example of this type of game, we can study the behavior of systems resembling multi-agent human decision-making. Just as humans may reduce computational complexity by mapping game instances to familiar game "templates", we may discover that such a strategy, as well as other strategies that are imperfect yet reasonable, provide benefits to engineered systems of agents.

Background. Economists and sociologists have proposed many variants and contexts of the Ultimatum Game that seek to address the divergence between the "rational" Nash equilibrium strategy and observed human behavior, for example, examining the game when played in different cultures, with members of different communities, where individuals are replaced by groups, where the players are autistic, and when one of the players is a computer. Interestingly, isolated non-industrialized cultures, people who have studied economics, groups, autists, and playing with a computer all tend to lead to less cooperative behavior [6, 4, 3, 1, 2]. Neuro-economists have suggested that the study of strategic interactions must be done with consideration given to the neurological constraints imposed by human biology [8, 7]. Evolutionary game theorists have examined replicator dynamics and other adaptive dynamics (trial and error, imitation, inheritance), which all tend to converge to the self-interested Nash equilibrium. By adding the ability for the proposer to retrieve information about the recipient's past actions, Nowak et al. show that these dynamics can converge to fair play [5]. However, in this work, proposers are paired with recipients randomly and receive this information without cost, both somewhat unrealistic assumptions. In this paper, we treat the selection of a potential recipient as part of the proposer's strategy, and reputation (or more generally, beliefs about other players' strategies) is uncovered naturally through the interactions allowed by the game. Thus, knowledge is gained only through an exploration process, mirroring what is often a real-world truth.

2 The Social Ultimatum Game

The Ultimatum Game, is a two-player game where a player, P_1 proposes a split of an endowment $e \in \mathbb{N}$ to another player P_2 where P_2 would receive $q \in \{0, \delta, 2\delta, \ldots, e-\delta, e\}$ for some value $\delta \in \mathbb{N}$. If P_2 accepts the offer, they receive q and P_1 receives $e - q$. If P_2 rejects, neither player receives anything. The subgame-perfect Nash equilibrium states that P_1 offer $q = \delta$, and P_2 accept. This is because a "rational" P_2 should accept any offer of $q > 0$, and P_1 knows this. Yet, humans make offers that exceed δ, even making "fair" offers of $e/2$, and reject offers greater than the minimum.

To represent the characteristics that people operate in societies of multiple agents and repeated interactions, we introduce the Social Ultimatum Game. There are N players, denoted $\{P_1, P_2, \ldots, P_N\}$, playing K rounds, where $N \geq 3$. The requirement of having at least three players in necessary to give each player a choice of whom to interact with.

In each round k, every player P_m chooses a single potential partner P_n and makes an offer $q^k_{m,n}$. Each player P_n then considers the offers they have received and makes a decision $d^k_{m,n} \in \{0, 1\}$ with respect to each offer $q^k_{m,n}$ to either accept (1) or reject (0) it. If the offer is accepted by P_n, P_m receives $e - q^k_{m,n}$ and P_n receives $q^k_{m,n}$, where e is the endowment to be shared. If an offer is rejected by P_n, then both players receive 0 for that particular offer in round k. Thus, P_m's reward in round k is the sum of the offers they accept from other players (if any are made to them) and their portion of the proposal they make to another player, if accepted:

$$r^k_m = (e - q^k_{m,n})d^k_{m,n} + \sum_{j=1\ldots N, j\neq m} q^k_{j,m} d^k_{j,m} \tag{1}$$

The total rewards for P_m over the game is the sum of per-round winnings, $r_m = \sum_{k=1}^{K} r^k_m$.

3 Adaptive Agents Model

In order to create mathematical models of human play for the Social Ultimatum Game that can yield results that match observed phenomena, we need to incorporate some axioms of human behavior that may be considered "irrational". The desiderata include assumptions that:

- People start with some notion of a fair offer
- People will adapt these notions over time at various rates based upon their interactions
- People have models of other agents
- People will choose the best option while occasionally exploring for better deals

While these rules certainly do not include all human characteristics, in this paper, we investigate the behaviors that can emerge from a mathematical model based solely on these axioms.

3.1 Characterizing the Players

Each player P_m is characterized by three parameters:

- α_m^k : Player m's acceptance threshold at time k
- β_m : Player m's reactivity
- γ_m : Player m's exploration likelihood

The value of $\alpha_m^k \in [0, e]$ is P_m's notion of what constitutes a "fair" offer at time k and is used to determine whether an offer to P_m, i.e., $q_{n,m}^k$, is accepted or rejected. The value of $\beta_m \in [0, 1]$ determines how quickly the player will adapt to information during the game, where zero indicates a player who will not change anything from their initial beliefs and one indicates a player who will solely use the last data point. The value of $\gamma_m \in [0, 1]$ indicates how much a player will deviate from their "best" play in order to discover new opportunities where zero indicates a player who never deviates and one indicates a player who always does. Each player P_m keeps a model of other players in order to determine which player to make an offer to, and how much that offer should be. The model is composed of the following values:

- $a_{m,n}^k$: P_m's estimate of P_n's acceptance threshold
- $\bar{a}_{m,n}^k$: Upper bound on $a_{m,n}^k$
- $\underline{a}_{m,n}^k$: Lower bound on $a_{m,n}^k$

Thus, P_m has a collection of models for all other players $\{[\underline{a}_{m,n}^k a_{m,n}^k \bar{a}_{m,n}^k]\}_n$ for each round k. The value $a_{m,n}$ is the P_m's estimate about the value of P_n's acceptance threshold, while $\underline{a}_{m,n}^k$ and $\bar{a}_{m,n}^k$ represent the interval of uncertainty over which the estimate could exist.

3.2 Adaptation Rules

During the course of the game, each player will engage in a variety of actions and updates to their models of agents. Below, we present our model of how our adaptive agents address those actions and model updates. For simplicity, we will assume that $\delta = 1$.

Making Offers: In each round k, P_m may choose to make the best known offer, denoted $\tilde{q}_m^k$, or explore to find someone that may accept a lower offer. If there are no gains to be made from exploring, i.e., the best offer is the minimum offer ($\tilde{q}_m^k = \delta = 1$), a player will not explore. However, if there are gains to be made from exploring, with probability γ_m, P_m chooses a target P_n at random and offers them $q_{m,n}^k = \tilde{q}_m^k - 1$. With probability $1 - \gamma_m$, P_m will choose to exploit. We introduce two approaches by which P_m can choose their target. In both cases, the target is chosen from the players who have the lowest value for offers they would accept, and the offer is that value:

$$q_{m,n}^k = \lceil a_{m,n}^k - \epsilon \rceil \text{ where } n \in \arg\min_{\tilde{n} \neq m} \lceil a_{m,\tilde{n}}^k \rceil \tag{2}$$

where $0 < \epsilon < \delta$ is some constant threshold. The previous equation characterizes an equivalence class of players from which P_m can choose a target agent. The ϵ parameter is used to counter boundary effects in the threshold update, discussed below. The difference in the approaches are as follows.

- The target agent from the equivalence class is chosen using *proportional reciprocity*, by assigning likelihoods to each agent with respect offers made in some history window.
- The target agent is chosen uniformly over all agents in the equivalence class.

Accepting Offers: For each offer $q_{m,n}^k$, the receiving player P_n has to make a decision $d_{m,n}^k \in \{0, 1\}$ to accept or reject it. The acceptance rule checks if the offer exceeds their threshold:

$$\text{If } q_{m,n}^k \geq \lceil \alpha_m^k - \epsilon \rceil, \text{ then } d_{m,n}^k = 1, \text{ else } d_{m,n}^k = 0 \tag{3}$$

Updating Acceptance Threshold: The acceptance threshold is a characterization of what the agent considers a "fair" offer. Once an agent is embedded within a community of players, the agent may change what they consider a “fair” offer based on what type of offers they are receiving. We model this adaption using a convex combination of the current threshold and the received offers. The rate of adaptation is determined by the player’s adaptivity parameter denoted β_m. Let the set of received offers be defined as: $R_m^k = \{q_{i,m}^k : i \neq m, q_{i,m}^k > 0\}$. If $|R_m^k| \geq 1$, then $\alpha_m^{k+1} =$

$$(1-\beta_m)^{|R_m^k|}\alpha_m^k + \frac{(1-((1-\beta_m)^{|R_m^k|})}{|R_m^k|}\sum_i q_{i,m}^k \tag{4}$$

If $|R_m^k| = 0$, then $\alpha_m^{k+1} = \alpha_m^k$. Thus, offers higher than a player’s expectation will raise its expectation and offers lower than its expectation will lower it.

Updating Threshold Estimate Bounds: As a player P_m makes an offer $q_{m,n}^k$ and receives feedback (a decision) on the offer $d_{m,n}^k$, it gains information about P_n’s acceptance threshold. We may update the bounds on the estimates of P_n’s threshold using the following set of rules:

If P_n rejects P_m’s offer, then the lower bound for the acceptance threshold must be at least the offer that was rejected:

$$q_{m,n}^k > 0, d_{m,n}^k = 0 \ \Rightarrow\ \underline{a}_{m,n}^{k+1} = \max\{q_{m,n}^k, \underline{a}_{m,n}^k\} \tag{5}$$

If P_n accepts P_m’s offer then the upper bound for the acceptance threshold for that player must be at most the offer that was accepted:

$$q_{m,n}^k > 0, d_{m,n}^k = 1 \ \Rightarrow\ \bar{a}_{m,n}^{k+1} = \min\{q_{m,n}^k, \bar{a}_{m,n}^k\} \tag{6}$$

The next two conditions occur because acceptance thresholds are dynamic and the bounds for estimates on thresholds for other players may become inaccurate and may need to be reset. If P_n rejects P_m’s offer and that offer was at least P_m’s current upper bound, then the upper bound increases to the “fair” offer that P_m expects P_n will accept:

$$q_{m,n}^k > 0, d_{m,n}^k = 0, q_{m,n}^k \geq \bar{a}_{m,n}^k \ \Rightarrow\ \bar{a}_{m,n}^{k+1} = \lceil e/2 \rceil \tag{7}$$

If P_n accepts P_m’s offer, however the offer is lower than P_m’s current lower bound estimate for P_n, then decrease the lower bound to zero:

$$q_{m,n}^k > 0, d_{m,n}^k = 1, q_{m,n}^k \leq \underline{a}_{m,n}^{k+1} \ \Rightarrow\ \underline{a}_{m,n}^{k+1} = 0 \tag{8}$$

Updating Threshold Estimates: The following set of rules are used to modify a player’s estimates of the other players thresholds: When P_n accepts P_m’s offer, P_m’s estimate of their acceptance threshold becomes closer to the lower bound, while upon rejection the estimate moves closer to the upper bound. Both of these updates are done using a convex combination of the current value and the appropriate bound as follows:

$$d_{m,n}^k = 1 \ \Rightarrow\ \quad a_{m,n}^{k+1} = \min\{\beta_m\, \underline{a}_{m,n}^{k+1} + (1-\beta_m)a_{m,n}^k, \bar{2}a_{m,n}^{k+1}\} \tag{9}$$

$$d_{m,n}^k = 0 \ \Rightarrow\ \quad a_{m,n}^{k+1} = \max\{\beta_m\, \bar{a}_{m,n}^{k+1} + (1-\beta_m)a_{m,n}^k, \underline{a}_{m,n}^{k+1} + 2\epsilon\} \tag{10}$$

The *min* and *max* operators ensure that players don’t make unintuitive offers (such as repeating a just rejected offer) when the adaptation rate is not sufficiently high. The adaptive agent described above fulfills the properties of the desiderata prescribed to generate behavior that is more aligned with our expectations in reality.

4 Experiments

Stability of Fair Players: First, we investigated a population of fair agents and discover the conditions under which they maintain their initial characteristics. We choose scenarios where a population of 6 agents of the same type play a game with e=10 for 100 rounds. Each agent begins with an acceptance threshold of $\alpha_m^0 = 5$, and estimates of other players’ thresholds at $a_{m,n}^0 = 5$, with the thresholds bounds at $\bar{a}_{m,n}^0 = 5$ and $\underline{a}_{m,n}^0 = 4$, and $\epsilon = 0.1$. We run 40,000 instances of the game where $\beta_m \in \{0.00, 0.05, 0.01, \ldots, 0.5\}$ and $\gamma_m \in \{0.00, 0.01, 0.02 \ldots, 0.8\}$.

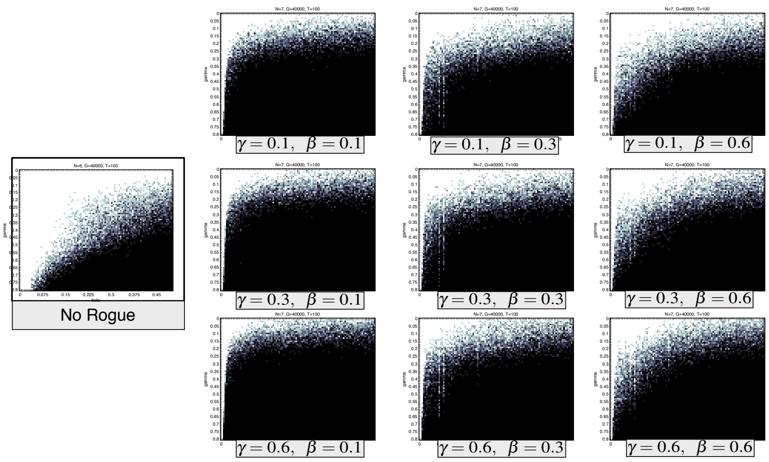

Figure 1: Stability of 6 “fair” players with and without an additional rogue agent. The heat map on the left depicts stability when there is no rogue. The x and y axes denotes the adaptivity (β) and the exploration (γ) parameters of the fair players, and darker squares in the heat map indicate higher probabilities that at least one agent will eventually accept an “unfair” offer. The nine figures on the right depict the stability when an additional rogue player with $\beta_m, \gamma_m \in \{0.1, 0.3, 0.6\}$ is added.

For each game type, characterized by the β and γ values of the agents, we record the likelihood that the acceptance threshold of at least one of the agents falls below 4.1, indicating that some player would accept “unfair” offers of 4. The players use the proportional reciprocity method for selecting their target agents. The results are shown in the heat map in Figure 1 (left). The x-axis denotes the adaptivity (β) values increasing from left to right and the y-axis denotes the exploration (γ) values increasing from top to bottom. Darker squares indicate higher probabilities that an agent accepts an offer of 4.

We see that at low adaptation and exploration rates, the society is able to maintain their egalitarian nature, but as both rates increase they start falling away. We investigated the scenario for horizons of 1000 rounds at a coarser discretization and the characteristics of the results were maintained.

The Effect of a Rogue Agent: To further investigate the stability of an egalitarian society, we observed the effect of adding a player who does not start with the societal norms. Here, we take the scenarios described earlier and add an agent with an acceptance threshold of $\alpha_m^0 = 1$, and estimates of other players' thresholds at $a_{m,n}^0 = 1$, with the thresholds bounds at $\bar{a}_{m,n}^0 = 5$ and $\underline{a}_{m,n}^0 = 4$, and $\epsilon = 0.1$. We note that the rogue agent's initial acceptance threshold estimate is lower than the lower bound of the estimate.

The intended effect of this is that the rogue will initially offer $q_{m,n}^k = 1$ once to each agent and then move its estimate up to $q_{m,n}^k = 1$ and then adapt within the bounds based on its adaptivity rate. We investigate 9 instances of rogue agents for $\beta_m, \gamma_m \in \{0.1, 0.3, 0.6\}$. The resulting heat maps are shown in Figure 1. Darker squares indicate higher probabilities that an agent starts accepting offers of 4. The square on the far left indicates the stability when there is no rogue agent.

The presence of a single rogue agent who only makes a single “lowball” offer and then small “undercutting” offers can significantly affect the stability of the population remaining egalitarian. As expected, lesser adaptivity and greater exploration rates of the rogue implied greater impact on stability. Adaptivity, or lack thereof, seems to have a greater influence than the exploration rate.

Reciprocity vs. Randomness: We look at two societies of agents where $N = 5$ and $N = 6$. All agents are the egalitarian type described earlier with $\gamma = 0.1$ and $\beta = 0.3$. The results for individual traces of a single game of each type when the algorithm for making offers used proportional reciprocity and when agents used random selection are shown in Figure 2.

In the figure, each heat map reflects the number of times a row player made an offer to a column player. A darker color represents a larger number of offers made from that row player to that column

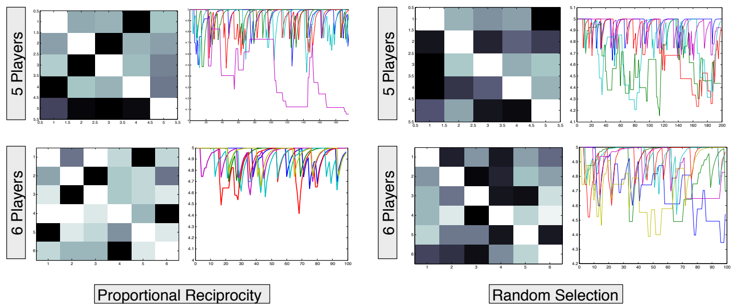

Figure 2: Reciprocity vs. Randomness for 5 and 6 "Fair" Players. A detailed description of this figure is given in the text.

player (The white diagonal indicated no player made an offer to itself). The subfigure to the right of each heat map shows the evolution of the acceptance thresholds of all the agents over time. We see that for random selection, the heat map of offers has no obvious pattern. On the other hand, under proportional reciprocity, agents seem to have paired up to form sustained partnerships.

In the 5-player game with reciprocity (top, left) P_1-P_4 is a partnership as is P_2-P_3, while P_5 is left out of the partnerships. We also see that P_5 makes offers uniformly to the other agents but does not receive many offers from them. Its acceptance threshold (the purple line) dives down due to it being a recipient solely of exploration offers and occasional offers due to proportional reciprocity. In the 6-player game with reciprocity (bottom, left), we have the following partnerships: P_2-P_3, P_1-P_5 and P_4-P_6. The 6-player game with reciprocity has much higher stability than all the other games as the acceptances threshold for all agents do not vary far from 5.

The games with random selection have multiple agents whose thresholds dive as the random nature of the offers let various agents be periodically ignored over the course of the trace. The key here is that reciprocity leads to the evolution of stable partnerships that help keep the egalitarian society stable as long as members are not ignored.

These results, and the overall issues explored in this paper, provide us with a framework for upcoming experiments with human subjects.

Acknowledgements. Yu-Han Chang is supported by a grant from the AFOSR MURI program, and Tomer Levinboim is supported by a USC Annenberg Graduate Fellowship.

References

[1] C. R. P. J. Carnevale. Group choice in ultimatum bargaining. *Organizational Behavior and Human Decision Processes*, 72(2):256–279, 1997.

[2] R. H. Frank, T. Gilovich, and D. T. Regan. Does studying economics inhibit cooperation? *The Journal of Economic Perspectives*, 7(2):159–171, 1993.

[3] E. Hill and D. Sally. Dilemmas and bargains: Theory of mind, cooperation and fairness. Working paper, University College, London, 2002.

[4] A. G. S. Mascha van't Wout, René S. Kahn and A. Aleman. Affective state and decision-making in the ultimatum game. *Experimental Brain Research*, 169(4):564–568, 2006.

[5] M. A. Nowak, K. M. Page, and K. Sigmund. Fairness versus reason in the ultimatum game. *Science*, 289(5485):1773 – 1775, September 2000.

[6] H. Oosterbeek, R. Sloof, and G. van de Kuilen. Differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics*, 7:171–188, 2004.

[7] A. G. Sanfey. Social decision-making: Insights from game theory and neuroscience. *Science*, 318(5850):598 – 602, October 2007.

[8] A. G. Sanfey, J. K. Rilling, J. A. Aronson, L. E. Nystrom, and J. D. Cohen. The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626):1755–1758, 2002.