Towards Cooperating in Repeated Interactions Without Repeating Structure

Huy Pham
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Physical Sciences and Mathematics Commons

BYU ScholarsArchive Citation
Pham, Huy, "Towards Cooperating in Repeated Interactions Without Repeating Structure" (2020). Theses and Dissertations. 8473.
https://scholarsarchive.byu.edu/etd/8473

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Towards Cooperating in Repeated Interactions Without Repeating Structure

Huy Pham

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Science

Jacob Crandall, Chair
Michael Goodrich
Kent Seamons

Department of Computer Science
Brigham Young University

Copyright © 2020 Huy Pham
All Rights Reserved
ABSTRACT

Towards Cooperating in Repeated Interactions Without Repeating Structure

Huy Pham
Department of Computer Science, BYU
Master of Science

A big challenge in artificial intelligence (AI) is creating autonomous agents that can interact well with other agents over extended periods of time. Most previously developed algorithms have been designed in the context of Repeated Games, environments in which the agents interact in the same scenario repeatedly. However, in most real-world interactions, relationships between people and autonomous agents consist of sequences of distinct encounters with different incentives and payoff structures. Therefore, in this thesis, we consider Interaction Games, which model interactions in which the scenario changes from encounter to encounter, often in ways that are unanticipated by the players. For example, in Interaction Games, the magnitude of payoffs as well as the structure of these payoffs can differ across encounters. Unfortunately, while there have been many algorithms developed for Repeated Games, there are no known algorithms for playing Interaction Games. Thus, we have developed two different algorithms, augmented Fictitious Play (aFP) and augmented S# (Aug-S#), for playing these games. These algorithms are designed to generalize Fictitious Play and S# algorithms, which were previously created for Repeated Games, to the more general kinds of scenarios modeled by Interaction Games. This thesis primarily focuses on the evaluation of these algorithms. We first analyze the behavioral and performance properties of these algorithms when associating with other autonomous algorithms. We then report on the results of a user study in which these algorithms were paired with people in two different Interaction Games. Our results show that while the generalized algorithms demonstrate many of the same properties in Interaction Games as they do in Repeated Games, the complexity of Interaction Games appear to alter the kinds of behaviors that are successful, particularly in environments in which communication between players is not possible.

Keywords: repeated interactions, game theory, machine learning
# Table of Contents

List of Figures vi

List of Tables x

1 Introduction 1

1.1 Thesis Statement .................................................... 3

2 Related Work 4

2.1 Relationship to Other Games ......................................... 4

2.1.1 Repeated Games (RGs) .............................................. 4

2.1.2 Stochastic Games .................................................. 4

2.2 Solution Concepts for Interaction Games .......................... 5

2.2.1 Maximin Strategy .................................................. 5

2.2.2 The Attack Strategy ............................................... 6

2.2.3 Nash Bargaining Solution ........................................ 6

2.3 Algorithms for Playing Games ....................................... 7

2.3.1 Fictitious Play .................................................... 7

2.3.2 Trigger Strategies ................................................ 8

2.3.3 S# ................................................................. 8

3 Augmented Algorithms for Interaction Games 10

3.1 Trigger Strategies ................................................... 10

3.1.1 Offers ............................................................. 11
7 Conclusion and Future Work 49

Appendices 52

A Questionnaire 53

B The Games 59

References 62
List of Figures

1.1 Example IG with four rounds (stage games). In each round, the players play a matrix game wherein the row player selects the row and the column player the column. The resulting joint action identifies a cell in the matrix assigning payoffs to the row and column players, respectively. 2

3.1 Payoff solutions for a given IGs. Black points represent all possible payoff solutions for the IGs. Blue dashed lines is maximin values for each player. The green dot is the Nash bargaining solution. The red points are solutions computed by the offers for different values of rho. 11

3.2 Rankings created by six high-level strategies σ for two different stage games. Precise definitions of the high-level strategies are provided in Appendix A. 13

4.1 Example normal-form games from each payoff family as classified by Bruns [3]. The highlighted solutions are the pure-strategy one-shot NE of each game. 20

4.2 Average payoffs achieved by aFP, Fair (ρ = 0) and Bully (ρ = 0.5) in self play in IGs with stage games drawn from Bruns’s payoff families [3]. Global NBS is the NBS of all the possible payoff combinations throughout the IG. Stage NBS is the average NBS value of all stage game. 23
4.3 (Left) Average payoffs obtained per round by aFP in self play in IGs of various types. aFP’s performance is compared with the average values from the stage games of (1) the maximin value, (2) the lowest NE, (3) the highest NE, and (4) the Nash bargaining solution [15] of the stage games. Results are averaged over 50 IGs each. Error bars show the standard error of the mean. (Right) The percentage of rounds that aFP played a best response in self to the action played by its partner.

4.4 Percentage of rounds in which aFP played a best response to its partner’s action in self play. Error bars (although negligible) show the standard error of the mean.

4.5 Aug-S#'s disappointment as a percentage of maximum payoff vs. different agents in mixed games.

4.6 The average payoff (as a percentage of the NBS) obtained by Aug-S# in self play in the (left) Unfair and (right) Mixed games shown in the Appendix B.

5.1 Two rounds in an unfair game. (Left) The payoff matrix in the first round. (Right) The payoff matrix in the second round.

5.2 User Interface provided to participants when cheap talk before each round was enabled.

5.3 User Interface provided to participants when when cheap talk was disabled.

5.4 Predefined speech acts participants could include in their messages in each round.

6.1 The average percentage of rounds in which the players played a best response when paired with people in unfair game (left) with cheap talk and (right) without cheap talk. The dots represent the average best response percentages of the player types and the lines are error bars. The error bars reflect 95% confident intervals.
6.2 The average percentage of rounds in which the players played a best response when paired with people in mixed game (left) with cheap talk and (right) without cheap talk. The dots represent the average best response percentages of the player types and the lines are error bars. The error bars reflect 95% confidential intervals.

6.3 Tukey pairwise comparisons showing the average percentage of rounds in which the players played a best response when paired with people over all games and conditions. The dots represent the average best response percentages of the player types and the lines are error bars. The error bars reflect 95% confidential intervals.

6.4 aFP’s best response percentage through 50 rounds in self-play (left) in unfair game and (right) in mixed game.

6.5 Average payoffs as a percentage of the NBS over both game (Left) with cheap talk and (Right) without cheap talk. The error bars show the standard errors.

6.6 Payoffs for individual interactions when cheap talk was permitted.

6.7 Payoffs for individual interactions when cheap talk was not permitted.

6.8 Average payoffs as NBS percentage in self-play in unfair game (Left) with cheap talk and (Right) without cheap talk. The error bars show the standard errors.

6.9 Average payoffs as NBS percentage in self-play in mixed game (Left) with cheap talk and (Right) without cheap talk. The error bars show the standard errors.

A.1 Demographic Questionnaire

A.2 Post-Experiment Questionnaire

B.1 The Unfair Game used in the user study. Each round consisted of a normal-form game drawn from the Unfair payoff family [3].
B.2 The Mixed Game used in the user study. Each round consisted of a normal-form
game randomly drawn from among the Bruns’ seven payoff families [3]. . . . 61
List of Tables

4.1 Expected behavior attributes of generalized algorithms in IGs ............... 19

5.1 User Study Schedule .................................................. 30

5.2 Users’ majors information .......................................... 31

7.1 A summary of the performance and behavioral attributes of our augmented
    algorithms in IGs as demonstrated by the results presented in this thesis. . . 51
Chapter 1

Introduction

As humans, we frequently interact with the same people over extended periods of time, sometimes on a daily basis. These individual encounters are the building blocks of human relationships. Social exchange theory [12] poses that every person gains and loses something with every interaction. Researchers have tried to formalize these interactions through different models. Games are one of the well-studied models of these interactions.

Many different game types have been created to model repeated interactions between people and other intelligent entities. Repeated games and their variants, which have been generalized as stochastic games [18], are perhaps the most frequently studied of these models. These types of games can be used to study the long-term relationship between people and machines. However, we find that no game type conveys the situation we want to study. In real life, we do not always encounter the same situation repeatedly. For example, unanticipated encounters occur frequently in many relationships. To have a thorough predictive model of the scenarios likely to be encountered through the interaction may be costly and even impossible. Thus, in this thesis, we present the concept of Interaction Games (IGs), which are stochastic games in which the future game stages or rounds of interaction are unknown and can change over time.

An interaction game (IG) is defined by a set of players $I$, who interact with each other in a sequence of games $G = (g_1, g_2, \cdots, g_T)$. Here, $g_t$ defines the $t^{th}$ encounter of the players, and $T$ defines the number of rounds in the IG. Each game $g_t$ can be of any finite game form, including a normal-form (matrix) game, an extensive-form game, or a finite stochastic game.
Stage game $g_1$  | Stage game $g_2$  | Stage game $g_3$  | Stage game $g_4$
---|---|---|---
| $c$ | $d$ | $e$ | $f$ | $g$ | $h$ | $i$
A | 0, 0 | 7, 7 | 20, 4 | 3, 4 | 0, 6 | 1, 8 | 1, 10
B | 14, 4 | 2, 1 | 9, 3 | 5, 0 | 2, 1 | 0, 0 | 3, 1
C | 8, 10 | 6, 5 | 8, 4 | | | |

Figure 1.1: Example IG with four rounds (stage games). In each round, the players play a matrix game wherein the row player selects the row and the column player the column. The resulting joint action identifies a cell in the matrix assigning payoffs to the row and column players, respectively.

Regardless of game form, the outcome of $g_t$ is a vector of finite payoffs obtained by the players, denoted by $r^t = (r^t_1, \ldots, r^t_I)$, where $r^t_j$ is the payoff to player $j$. For simplicity, we assume that rewards are translated such that $r^t_j \geq 0$, and all players rewards are scaled equivalently.

An illustrative IG of short duration is shown in Figure 1.1. The game illustrates that players of IGs must deal with encounters having different equilibrium characteristics, symmetry characteristics, and importance. Given that players do not know future stage games, these characteristics make efficient cooperation in such interactions challenging.

The set of possible IGs is vast. Thus, as a starting point to studying IGs in this thesis, we make several stipulations. First, we assume that each game $g_t$ is a two-player normal-form game. Let $A(g_t) = A_i(g_t) \times A_{-i}(g_t)$ be the set of joint actions in stage game $g_t$, where $A_j(g_t)$ is the action set of player $j$ in game $g_t$. Let $r^t_j(a)$ denote the reward received by player $j$ when the joint action $a = (a_i, a_{-i})$, where $a_i \in A_i(g_t)$ and $a_{-i} \in A_{-i}(g_t)$, is played in game $g_t$. Second, in each round, we assume that players know $g_t$, and that they can observe the actions taken by each other. Third, we assume that $T$ is unknown to both players and is likely large. Fourth, we assume that the choices made by the players in game $g_t$ have little or no impact on subsequent stage games ($g_{t+1}$ through $g_T$). Finally, at time $t$, $g_{\tau}$ is unknown to the players for all $\tau > t$.

Because of the newness of IGs, there is no known algorithm capable of achieving desirable results in this type of game. However, over the years, scientists have developed
many algorithms to increase the ability of Artificial Intelligence (AI) to interact with other agents, including humans, in many other game settings. Thus, in this thesis, we will study how well these approaches can be extended to IGs.

Specifically, in this thesis, we extended two known algorithms to IGs (Fictitious Play [2] and S# [5]). We then evaluated the performance of these algorithms both via simulation and via a user study. Our results showed that at least one of our algorithms was able to interact effectively against other agents as well as humans in each setting we considered. Compared to humans, our algorithms achieved similar results when interacting with another human in this setting. Our extended version of S# performed as well as humans when communication was allowed. On the other hand, our extended version of FP achieved similar payoffs to that of humans when there was no communication. This research forms a basis for future research in repeated interactions in which the scenario in which the players interact changes over time.

1.1 Thesis Statement

In this research, we are interested in constructing agents that can interact with other intelligent agents in long-term interactions. We model these interactions as interaction games (IGs). We generalize FP [2], Trigger Strategies [11], and S# [5], algorithms previously created for repeated games (RGs), to IGs. We focus on evaluating and validating these algorithms via simulations and a user study. We are interested in determining whether our extended algorithms maintain the same performance attributes in IGs as the original algorithms have in RGs. In particular, we hypothesize that the empirical distribution of actions played by aFP (extended version of FP) in self play will converge to a Nash equilibrium. Furthermore, we hypothesize that Aug-S# (our extended version of S#) will, like the original algorithm, both (1) tend to minimize disappointment and (2) achieve cooperative solutions when associating with people and other algorithms that are willing to cooperate in IGs.
Chapter 2

Related Work

In this proposed research, we will design and analyze algorithms in IGs. IGs are designed to comply with our goal of studying and handling interactions in a series of time-varying, perhaps unexpected, encounters. In this section, we discuss how IGs relate to other game forms designed to model extended interactions between players. We also review algorithms that have been developed to solve similar types of games as well as the game theory related to the subject.

2.1 Relationship to Other Games

Interaction games are closely related to other commonly studied games, including repeated games and stochastic games.

2.1.1 Repeated Games (RGs)

IGs generalize RGs. An RG is an IG in which, for all $t$, $g_{t+1} = g_t$. In words, in each round of an RG, the players play the exact same game. Many algorithms have been developed for playing RGs. However, because the payoff matrix can change from round to round in IGs, these algorithms may not extend in a straight-forward fashion to IGs.

2.1.2 Stochastic Games

A stochastic game (SG) [18] is a tuple $(S, A, T, R)$, where

- $S$ is a set of states
• \(A = A_1 \times A_2 \times \ldots \times A_n\) (\(A_i\) is a set of actions for player \(i\))

• \(T\) is a transition function

• \(R\) is a reward or payoff function

In the case of two-player SG, the state transition is based on probabilities \(\pi(g_{t+1}|g_t, a_i, a_{-i})\) in which \(a_i\) and \(a_{-i}\) are actions of the players in round \(g_t\). SGs [17] can be used to model many interactions in the real world. Nevertheless, the world is sufficiently unpredictable that it is difficult to accurately assess the probability transitions between states. Therefore, IGs were designed as a form of SGs in which unpredictable or unknown future events produce a new game \(g_{t+1}\). Therefore, unlike in regular SGs, the new payoff structure \(g_t\) is generated prior to each round of an IG that is not based on some previously known probability distribution function.

2.2 Solution Concepts for Interaction Games

In this section, we will discuss several solution concepts for other related game types to IGs. These solution concepts have been used and studied repeatedly by different researchers for these types of games. These concepts will be the basic ideas in building algorithms for IGs.

2.2.1 Maximin Strategy

In repeated games, the maximin value [19] is the maximum expected payoff that a player can guarantee itself regardless of the action of its opponent. An agent should play this strategy when it believes its associate will act counter to its desires. The maximin value is formally defined as

\[
\nu_{i}^{\min}(g_t) = \max_{\pi_i \in \Pi_i(g_t)} \min_{\pi_{-i} \in \Pi_{-i}(g_t)} r_{i}^{g_t}(\pi_i, \pi_{-i})
\]  

(2.1)

where \(\Pi_j(g_t)\) is the set of legal probability distributions over player \(j\)’s action set in \(g_t\), and \(r_{i}^{g_t}(\pi_i, \pi_{-i})\) is the expected payoff to player \(i\) when the players play strategies \(\pi_i\) and \(\pi_{-i}\),
respectively. The player’s maximin strategy is the strategy it plays in this solution, given by:

\[
\pi_{mm}^i(g_t) = \arg \max_{\pi_i \in \Pi_i(g_t)} \min_{\pi_{-i} \in \Pi_{-i}(g_t)} r_{g_t}^i(\pi_i, \pi_{-i}) \tag{2.2}
\]

### 2.2.2 The Attack Strategy

When a player desires to punish its partner, it uses its attack strategy, or the strategy that minimizes its partner’s payoffs in a stage game. This strategy is given as follows:

\[
\pi_{\text{attack}}^i(g_t) = \arg \min_{\pi_i \in \Pi_i(g_t)} \max_{\pi_{-i} \in \Pi_{-i}(g_t)} r_{g_t}^i(\pi_i, \pi_{-i}) \tag{2.3}
\]

### 2.2.3 Nash Bargaining Solution

The goal of our algorithm is not to receive more reward than the agents which it interact with, but to maximize its own reward. To accomplish such a task, we have to clearly define a benchmark in which we can compare our result with. We use John Nash’s Bargaining Solution [15] as our desired result. The Nash Bargaining solution guarantees four axioms:

- Invariant to affine transformations or Invariant to equivalent utility representations
- Pareto optimality
- Independence of irrelevant alternatives
- Symmetry

In other words, Nash Bargaining Solutions \((v_{NBS}^i, v_{NBS}^{-i})\) are the results of the following optimization problem:

\[
\max_{r_t^i, r_{-i}^t} (r_t^i - v_{mm}^i)(r_{-i}^t - v_{mm}^{-i}) \tag{2.4}
\]

subject to: \((r_t^i, r_{-i}^t) > (v_{mm}^i, v_{mm}^{-i})\)
2.3 Algorithms for Playing Games

There have been many algorithms developed to solve different types of games. Nevertheless, repeated games have the closest relationship to interaction games. Hence, we will focus on the strategies developed to solve repeated games. For RGs itself, there are many different types of algorithms such as: tit-for-tat, Fictitious Play [10], multi-agent reinforcement learning (e.g., Minimax-Q [14], WoLF [1], M-Qubed [7], LOLA [9]), leader strategies (Bully, Godfather) [13], trigger strategies, etc. In this proposed research, we study how generalized version of three algorithms perform in IGs. These three algorithms are Fictitious Play, Trigger Strategies [11], and S# [5]. Thus, in this section, we briefly discuss these three algorithms.

2.3.1 Fictitious Play

One of the oldest and most studied algorithms for RGs is Fictitious Play (FP) [2, 10]. Despite its simplicity, FP performs relatively well in RGs. For example, in a recent simulation, FP ranked sixth among 25 algorithms [5]. FP, however, has a weakness. It tries to maximize its current payoff without considering the long-term effect of its action on the opponent. Therefore, FP often does not reach Pareto optimal payoffs.

The idea behind FP (Alg 1) is simple. The algorithm models its opponent based on the empirical distribution of the actions of its opponent. With the initial $\kappa^0(a_{-i}) = 1$, the estimated probability of each action for the opponent at round $t$ is:

$$
\gamma_t^{'}(a_{-i}) = \frac{\kappa_t^{'}(a_{-i})}{\sum_{a' \in A_{-i}(g)} \kappa_t^{'}(a')} 
$$

(2.5)

After that, the algorithm plays the best response to the most frequent or the highest probability actions. It is a greedy approach. FP then computes $\kappa_t+1^{'}(a_{-i})$ for each $a_{-i} \in A_{-i}(g)$ given $\kappa_t^{'}(a_{-i})$ and its observation of its partner’s action $a_{-i}^{'}$. The process is then repeated for the next round.
Algorithm 1 Fictitious Play for player $i$ for RGs.

**Input:** Game matrix $g$  
**Initialize:** $t \leftarrow 0$; $\kappa^0(a_{-i}) \leftarrow 1$ for all $a_{-i} \in A_{-i}(g)$  
**repeat**  
- For all $a_{-i} \in A_{-i}(g)$, compute $\gamma^t(a_{-i})$ using Eq. (2.5)  
- Select action $a^t_i \in br_i^g(\gamma^t)$  
- Observe partner’s action $a^t_{-i}$  
- For all $a_{-i} \in A_{-i}(g)$, compute  
  $$
  \kappa^{t+1}(a_{-i}) \leftarrow \begin{cases} 
  \kappa^t(a_{-i}) + 1 & \text{if } a_{-i} = a^t_{-i} \\
  \kappa^t(a_{-i}) & \text{otherwise}
  \end{cases}
  $$  
- $t \leftarrow t + 1$  
**until** Game Over

2.3.2 Trigger Strategies

A trigger strategy [11] for RGs is defined by two elements: an offer and a punishment. The offer is a proposal from one player to another in which each player will follow a particular pattern of actions throughout the game. If the other player does not follow the agreement, the agent then applies the punishment forever or until the game reaches a specific condition. The Folk Theorem states that pairs of trigger strategies can be used to generate an infinite number of Nash equilibria in many repeated games [11].

2.3.3 S#  

S# [5] and the communicating version of exploration-exploitation experts method (EEE [8]), EEE# [16] are two algorithms which have done well with repeated games. In a recent comparison, S# performed a little better than EEE# in playing repeated games. Therefore, we will take a deeper look into S#.

S# was designed for repeated games, and is an extension of the algorithm S++ [4]. S++ is a meta-algorithm that selects experts from a set of experts in an efficient way. Each expert in the set is a strategy for repeated games on its own such as maximin, trigger strategies,
etc. Later, the algorithm was expanded to perform successfully in repeated stochastic game [6].

S# is encoded with three components: aspiration level, potential, and communication. Its aspiration level is the agent’s payoff goal for each round. Potential is the expected payoff of each expert and is compared to the aspiration level. Communication gives S# the ability to talk to other agents. Because of the importance of these attributes, we will explain them more clearly below.

**Aspiration level** The algorithm starts with an aspiration level or a payoff goal. If the payoff on the current round satisfies the aspiration level, S++ will keep using the current expert. Otherwise, S++ would compile a new set of satisfying experts. The experts in the satisfying set will have potentials value higher than the aspiration values. S++ will then randomly select a new expert from this set to continue the game.

**Potential** Potential is the highest expected payoff that we can reasonably believe that the expert will receive when the opponent plays a best response against its strategy. Formally, a best response is the strategy that yields the player its expected payoff given the strategy of the other player. Thus, potential is defined as follows:

\[ z_i^t(\phi) = \mu_i(\phi, br_{-i}(\phi)), \]  

(2.7)

where \( br_{-i}(\phi) \) denotes player \(-i\)’s best response to the strategy \( \phi \) played by player \( i \).

**Communication** S# took its form when communication was added to S++. The algorithm gained the ability to converse with its opponent. This component helped S# cooperate with the other player by giving suggestion and intent of its future action in each round. This was to be the key feature in the improvement of S++. When the other players followed S#’s suggestion and cooperated, they received better payoff than they would have otherwise [5].
Chapter 3

Augmented Algorithms for Interaction Games

Since S# has been shown to be successful in interacting with people and other algorithms in Repeated Games [5], we generalize S# to interaction games and then evaluate the extent to which it can effectively interact with people in IGs as it does in RGs. As the first step towards generalization, we use strategies from other algorithms that will be used as expert strategies for S#.

There are many strategies used in RGs. Some of the most common strategies and solution concepts are maximin, minimax, and the best response. Besides these strategies, we introduce a new set of experts for Aug-S#: Fair Offer, Bully Offers, Bullied Offers, Pleaser, and Augmented Fictitious Play. However, we do not provide minute details for these algorithms since the primary contribution of this thesis is the evaluation (not the development) of these algorithms via simulation and user study.

3.1 Trigger Strategies

As discussed in the previous chapter, a trigger strategy consists of two parts: the offer and the punishment. However, unlike the previous definition of trigger strategies in which punishment continues forever once one deviates from the offer, our trigger strategies only punish until a certain condition is met. We briefly overview each part of the trigger strategies in this section.
Figure 3.1: Payoff solutions for a given IGs. Black points represent all possible payoff solutions for the IGs. Blue dashed lines is maximin values for each player. The green dot is the Nash bargaining solution. The red points are solutions computed by the offers for different values of rho.

3.1.1 Offers

The offer in our trigger strategies focused on moving the average payoff toward a certain point on the payoff convex hull using a $\rho$ value. Figure 3.1 shows a sample set of payoff pairs from a particular IG. According to the Folk Theorem [11], each payoff pair in which both players get a payoff higher than its maxmin value (dashed blue line in the figure) can be sustained as a Nash equilibrium of the repeated game. We seek to compute offers within this set that reside or or near the Pareto boundary. These offers are shown in red in the figure, and are parameterized by the value $\rho$. When $\rho = 0$, the algorithm computes an offer that tends to be near the Nash Bargaining Solution (the Fair Offer). $\rho$ values greater than 0 produce offers that favor the row player. We call these offers *Bully Offers*. On the hand, when $\rho < 0$, the offer favors the column player. We call these offers *Bullied Offers*. In this thesis, we named our trigger strategy agents in according to its characteristic. For example, if the $\rho$ value = 0.4, we named the agent bully4. On the other hand, if the $\rho$ value = −0.4, we named the agent bullied4.
3.1.2 Punishment

As in the trigger strategies defined for RGs, the punishment phase of our trigger strategies for IGs is achieved by playing the attack strategy (Eq. 2.3). The attack or minimax strategy must be played until the benefit obtained by player $-i$ for deviating from the offer is more than negated by the punishment. Formally, when player $-i$ deviates from the offer, it accrues guilt $g_{-i} = r_{-i}^t - \hat{r}_{-i}^t + \delta$, where $\hat{r}_{-i}^t$ is the payoff it should have received in the round had it conformed with the offer, and $\delta > 0$ is a constant value. In subsequent rounds, the guilt is updated as follows: $g_{-i} = g_{-i} + r_{-i}^t - \hat{r}_{-i}^t$. Once $g_{-i} \leq 0$, the punishment phase ends.

3.2 Pleaser

This algorithm is created as a counter to the trigger strategies. The algorithm assumes its partner is playing a trigger strategy, and simply tries to learn from experience the $\rho$ value of the offer in this trigger strategy. It then follows the actions prescribed to it in this offer. When communication is possible, the Pleaser strategy simply follows the solutions suggested by its partner when such suggestions are available.

3.3 Augmented Fictitious Play

In IGs, the payoff matrix can be different in each round. Hence, modeling the frequency of each action (the modeling mechanism used in FP) does not model the opponent very well. The same strategy can result in different actions because the payoff for each action pairs are not the same. Thus, rather than counting actions, our augmented version of FP (call augmented FP, or aFP) counts the number of times the player’s partner conforms with a high-level strategy that can be generalized across games. We used the seven high-level strategies listed in Figure 3.2.

Using this modification, aFP (Algorithm 2) then follows FP in a straight-forward fashion. First, the algorithm ranks each action in the game using each high-level strategy
from a set $\Sigma$ of high-level strategies. It then decides which high-level strategy best models the opponent’s strategy, given by

$$\sigma^* = \arg\max_{\sigma \in \Sigma} c^t_{\sigma}(1)$$

(3.1)

where $c^t_{\sigma}(k)$ is the number of times that $\sigma$’s $k$th-ranked action has been played up to round $t$. Finally, aFP will play the action associate with the best response strategy to the modeled opponent’s strategy.

To create action mappings across stages, each high-level strategy $\sigma \in \Sigma$ orders its partner’s actions in each round with respect to the quality of that action as measured by the utility function of the high-level strategy. Let $\Gamma_{g_t}(a_{-i})$ denote the rank of action $a_{-i}$ in stage game $g_t$ as determined by high-level strategy $\sigma$. The action orderings made by each $\sigma \in \Sigma$ equate actions across the stage games of the IG. For example, in the scenario presented in Figure 3.2, the high-level strategy Safe equates action $d$ in stage game $g_k$ with action $h$ in stage game $g_{k+1}$ since it gives those actions the same rank (i.e., $\Gamma_{Safe}^{g_k}(d) = \Gamma_{Safe}^{g_{k+1}}(h)$). On the other hand, the high-level strategy Max Social Welfare equates action $d$ with action $h$ because it orders the actions differently than Safe.

Recall that, in FP, $\kappa$ represents the number of times an action is played. For each $a_{-i} \in A_{-i}(g_t)$, we then have

$$\kappa^t(a_{-i}) = 1 + c^t_{\sigma^*}(\Gamma_{\sigma^*}^{g_t}(a_{-i}))$$

(3.2)
Algorithm 2 Augmented Fictitious Play (aFP) for player $i$.

**Input:** $G = (g_0, \cdots, g_T)$
**Initialize:** $t \leftarrow 0; c^0_\sigma(k) \leftarrow 0$ for all $k \in [1, N], \sigma$

repeat
  - For all $a_{-i} \in A_{-i}(g_t)$, compute $\kappa^t(a_{-i})$ using Eq. (3.3)
  - For all $a_{-i} \in A_{-i}(g_t)$, compute $\gamma^t(a_{-i})$ using Eq. (2.5)
  - Select action $a^t_i \in br^g_t(\gamma^t)$
  - Observe partner’s action $a^t_{-i}$
  - For all $\sigma$ and $a_{-i} \in A_{-i}(g_t)$, compute
    \[
    c^{t+1}_\sigma(x) \leftarrow \begin{cases} c^t_\sigma(x) + 1 & \text{if } x = a^t_{-i} \\ c^t_\sigma(x) & \text{otherwise} \end{cases} \tag{3.4}
    \]
  - $t \leftarrow t + 1$

until Game Over (when $t > T$)

where 1 is added to the quantity as a prior, so that $\kappa^0(a_{-i}) = 1$ (as in FP; see Algorithm 1).

In Eq. (3.2), we assumed that the number of actions available to player $-i$ is the same in each round. Since this is not always true, we can adjust the counts to represent how frequently they are played when available as follows:

\[
\kappa^t(a_{-i}) = 1 + c^t_\sigma(\Gamma^g_t(a_{-i})) \left( \frac{t - 1}{Z^t(\Gamma^g_t(a_{-i}))} \right) \tag{3.3}
\]

where $Z^t(\Gamma^g_t(a_{-i}))$ is the number of stage games up to round $t$ in which player $-i$ has had at least $\Gamma^g_t(a_{-i})$ actions in its action set.

When aFP is used as expert for Aug-S#, its potential is calculated as followed

\[
z^t_i(aFP) = \sum_{\tau=0}^t \max_{a \in A_i(g_t)} U_{g_t}(a) \tag{3.5}
\]

3.4 Augmented-S#

We begin by detailing the generalized version of S#, called Augmented-S# or Aug-S#, which is overviewed in Algorithm 3. Since many of the algorithmic mechanisms of S# rely on
the same game being played repeatedly by the players, we must modify these algorithmic mechanisms so that they can be used in IGs. In modifying these algorithmic mechanism, we must ensure that they maintain the important properties that make S# successful. Thus, in RGs, the Aug-S# algorithm behaves nearly identically to S# while also being able to implement the same principles of behavior in IGs.

Our generalization of S# is achieved by making changes to both S#’s mechanism for selecting which of its experts to follow and the experts themselves. We describe changes to each of these elements of the algorithm in turn in the following two subsection.

3.4.1 Selection Mechanism

Aug-S#’s selection mechanism is similar to the selection mechanism of the original algorithm, and operates as follows. First, it encodes (and learns) an aspiration level that the agent will aim to achieve each round. Second, Aug-S# calculates a potential for each expert. If an expert’s potential is higher than the aspiration level, then Aug-S# will consider selecting that expert (it is placed in the set of satisficing experts from with the algorithm selects which expert to follow). Third, the algorithm selects an expert from the set of satisficing experts. Fourth, the algorithm then follows this expert for \( \tau \) rounds. \( \tau \) is a parameter defined by the algorithm designer. Fifth, the algorithm calculates the ratio of the difference between the total aspiration and the total maximin value up to round \( t \) over the difference between the total NBS value and the total maximin value up to round \( t \). If this value is more than \( \epsilon \) less than that of the previous cycle, the algorithm selects a new expert from the set of satisficing experts using an \( e \)-greedy function (a departure from the original S++ algorithm for repeated games). Afterward, we repeat step four and five until the game is over.

Since the payoff matrices in IGs change every stage, we must modify the algorithm so that it can account for these changes in its selection mechanism. The following two changes are required so that algorithm maintains its same function in IGs as it uses in RGs:
1. Aspiration level – This is the value that Aug-S# will aim to achieve, on average, each round. The aspiration level should fulfill two characteristics: initially optimistic and converge to the average reward received in the long term (long-term accuracy). In the original algorithm, the initial aspiration level is typically set high, either at the highest payoff of the stage game or at the stage game’s Nash Bargaining Solution (NBS). It is then updated each round as a convex combination of the previous aspiration level and the reward received in the latest round. However, given that the NBS of each round of an IG is distinct, this mechanism does not guarantee that the aspiration level will be set optimistically, nor does it permit the aspiration level to be compared against values of an individual round (since each round is scaled differently). Thus, Aug-S# defines its aspiration level differently, but with the intent to still maintain the properties of (a) initial optimism and (b) long-term accuracy. Aug-S#’s aspiration level is defined as follows:

\[ \alpha_i(t) = \omega V_{i}^{NBS}(t) + (1 - \omega) \bar{R}(t), \]  

(3.6)

where is \( \bar{R}(t) \) the average payoff obtained by player \( i \) up to game \( t \), given by \( \bar{R}(t) = \frac{1}{t-1} \sum_{\tau=1}^{t-1} r_{i}^{\tau} \), and \( V_{i}^{NBS}(t) \) is the average NBS of each round up to time \( t \). \( \omega \) is decreased over time such that in early rounds \( V_{i}^{NBS}(t) \) dominates and in later rounds \( \bar{R} \) has more weight.

Algorithm, it is sometimes useful to use a normalized version of the aspiration level, which we define as

\[ \alpha_i'(t) = \frac{\alpha_i(t)}{V_{i}^{NBS}(t)}. \]  

(3.7)

2. Potential of experts – Aug-S# uses an estimate of the potentials of its experts to determine whether each expert could potentially produce a payoff that meets its aspiration level. Therefore, the potential of experts needs to be comparable to the aspiration level. Since the values of each round can vary substantially from round
to round in an IG, we must modify these potentials so that they are comparable to
the aspiration level as given in Eq. 3.6. To do this, we compute the potential of each
expert \( e_k \) as a summation of \( V_{i}^{e_k}(g_t) \) (the value of expert \( e_k \) in game \( g_t \) under ideal
circumstances) over all games played in the interaction so far. Formally,

\[
\rho_k(t) = \frac{1}{t-1} \sum_{j=1}^{t-1} V_{i}^{e_k}(g_j). \tag{3.8}
\]

Since \( \alpha(t) \) and \( \rho_k(t) \) are both computed over the same set of games, the two are
comparable. Hence, we can always compare an expert’s potential with the aspiration
level at any time \( t \).

3. Expert selection - After using an expert for \( \tau \) rounds, Aug-S# evaluates the effectiveness
of the expert. As in the original algorithm, Aug-S# continues to play its current expert
when its rewards meet its aspiration level. Otherwise, it selects a new expert from the
satisficing set of experts \( E(t) \) using an \( \varepsilon \)-greedy function. With probability \( r = \frac{1}{2 + \frac{t}{100}} \) is
used to decide if the method should choose greedily or just pick a random expert from
\( E(t) \). To choose greedily, the algorithm chooses its next expert from \( E(t) \) that has had
the highest productivity when used in the past. Here, we define productivity as:

\[
p_e(t) = \frac{R_e(t)}{V_{i}^{NBS}(t)}, \tag{3.9}
\]

where \( R_e(t) \) is the total payoff of expert \( e \) in the rounds in which it was used up to time
\( t \) and \( V_{i}^{NBS}(t) \) is the NBS value for player \( i \) in those rounds. Eq. (3.10) in Algorithm 3
summarizes this expert-selection mechanism.
Algorithm 3 Aug-S#

**Input:** Game $G$ and set of experts $E = \{e_1, \ldots, e_n\}$

**Initialize:** $t \leftarrow 1; \bar{r}(1) \leftarrow 0; \alpha \leftarrow V_i^{\text{NBS}}(1)$

**repeat**

Update potential $\rho_k(t)$ of each expert $e_k \in E$ (Eq. 3.8)

Compute $E(t) = \{e_k \in E : \rho_k(t) \geq \alpha(t)\}$

Select $e(t)$ from $E(t)$ using

$$
e(t) = \begin{cases} 
  e(t - \tau) & \text{if } \alpha'_i(t) \text{ is non-decreasing} \\
  \varepsilon\text{-greedy selection based on } p_e(t) & \text{otherwise}
\end{cases}
$$

(3.10)

Select action using $e(t)$ for $\tau$ rounds;

$t \leftarrow t + \tau$

Compute $\alpha(t)$ (Eq. 3.6)

**until** Game Over
Chapter 4

Performance and Behavioral Attributes of the Algorithms in Simulation

In this thesis, we seek to evaluate the extent to which the augmented algorithms overviewed in the previous chapter maintain the same characteristics in IGs that the original algorithms hold in RGs. A list of important attributes for these algorithms are show in Table 4.1. We evaluate whether our augmented algorithms maintain these properties via a set of simulations (reported in this chapter) and via a user study (reported in the subsequent two chapters).

For the simulation results presented in this chapter, we first describe a set of games in which we evaluate the algorithms. We then report the results of various simulations in the game environments to evaluate the performance attributes of each algorithm in IGs.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Expected Behavior</th>
</tr>
</thead>
</table>
| aFP        | - The empirical distribution of actions in self play converges to a NE. (P1)  
              - Learn to play best response against people. (P2) |
| Trigger Strategies | - Offers are NE of the repeated game and are near the Pareto boundary. (P3) |
| Aug-S#     | - Obtain low disappointment when associating with many other algorithms. (P4)  
              - Obtain payoffs near the NBS value in self-play. (P5)  
              - Achieve cooperative solutions when playing with people who are willing to cooperate. (P6) |

Table 4.1: Expected behavior attributes of generalized algorithms in IGs
Figure 4.1: Example normal-form games from each payoff family as classified by Bruns [3]. The highlighted solutions are the pure-strategy one-shot NE of each game.

4.1 Game Testbed

Since IGs have not been studied in prior work, we constructed a game testbed on which to evaluate algorithms. This game testbed consists of two forms of games. The first set of games are created by randomly generating a separate normal-form game for each round of the interaction (interactions of 100 rounds and 1000 rounds are considered). On the other hand, the second set of games are generated such that the players play a 2-action normal-form game in each round randomly generated to conform to a particular payoff family. These payoff families were categorized by Bruns with respect to the equilibrium characteristics of the game [3]. Thus, by testing the algorithms in games with each payoff family, we are able to observe how the payoff structures in and across interaction games impact the behaviors and performances of algorithms. A more detail explanation of these game families is presented below.

In the second set of games, each round’s game was generated by randomly selecting a payoff structure from the specified payoff family. In these payoff structures, numbers represent the ordinal preference ratings of the players, with 1 being the lowest preference and 4 being
the highest preference. Figure 4.1 shows example payoff structure from each payoff family, each of which has different equilibrium characteristics. Once the payoff structure is selected, four randomly generated numbers in the range 0 to 100 are chosen. We then replace the preferences in the payoff structure with the random number in increasing order, such that 1 would be replaced by the lowest generated number and so on. In this way, the players must solve a different game in each round with distinct scales and equilibrium attributes.

In this way, we generated eight different categories of games. In seven of these categories of games, each round’s game was generated from a payoff structure drawn from the same payoff family. In the eighth category, the game played in each round was randomly generated from all payoff structures.

A brief overview of each payoff family is as follows:

- **Win-win** games have a equilibrium solution that would give both players the highest possible payoff. For example, if the row player chooses to play A and the column player chooses to play Y, they would both receive 4 which is the highest payoff for both of them.

- **Biased** games have a equilibrium solution that would give one player the highest possible payoff while the other receives the second best payoff. For example, if the row player chooses to play B and the column player chooses to play Y, the row player would receive 4 and the column player only gains 3.

- **Second Best** games have a equilibrium solution that would give both players the second highest possible payoff. For example, if row player chooses to play B and column player chooses to play X, they would both receive 3 which is the second highest payoff for both of them.

- **Unfair** games have one equilibrium solution that would give a player the highest possible payoff while the other receives the third best payoff. For example, if the row
player chooses to play B and the column player chooses to play Y, the row player would receive 4 and the column player only gets 2.

- **Traps** games have an equilibrium solution which are Pareto Dominated by the solution in which both players play the other action. The equilibrium solution in our example is 3,2 which is Pareto dominated by 4,3 (the results when both players choose different actions).

- **Sad** games have an equilibrium solution which gives a player a second best payoff and the other the second worst payoff. In our example, when the row player chooses B and the column player chooses X. The row player receives 3 which is the second best payoff while the col players receive the second worst payoff of 2.

- **Cyclic** game type has no pure Nash Equilibrium solution. That means for any action pair, one of the players would want to move unilaterally to get a better payoff. In our example game, if row player chooses A and column chooses X, row player would want to change the action to B unilaterally to receive the better payoff of 2 instead of 1. That applies to one of the players for any solution in our example game.

50 games of each type were generated for both 100-round and 1000-round interactions, respectively. In short, the random games of different sizes (numbers of actions) are used to see how well the algorithms scale to more complex scenarios. On the other hand, more structured IGs based on Bruns’ payoff families are designed to study how well the algorithms adapt to different payoff structures.

### 4.2 Trigger Strategies

The qualities of the offers produced by the various trigger strategies are critical to the ability of Aug-S# to achieve cooperative solutions in self play and when associating with other people. As specified in Table 4.1, these offers should give both players greater than their maximin values. Ideally, they should produce payoffs on or near the Pareto frontier (Property
Figure 4.2: Average payoffs achieved by aFP, Fair ($\rho = 0$) and Bully ($\rho = 0.5$) in self play in IGs with stage games drawn from Bruns’s payoff families [3]. Global NBS is the NBS of all the possible payoff combinations throughout the IG. Stage NBS is the average NBS value of all stage game.

P3). Given that the quality of these offers are evaluated simultaneously with Aug-S#, we do not evaluate them in detail alone. However, Figure 3.1 illustrates that the algorithm achieves these characteristics in a single IG. Figure 4.2 also illustrates that the Fair offer ($\rho = 0.0$) produces payoffs essentially equivalent to the NBS value of the game across a broad range of games, while a Bully strategy $\rho = 0.5$ can potentially provide a player with an even higher payoff (contingent on the other player also conforming with this offer).

4.3 Augmented Fictitious Play (aFP)

We analyze the ability of aFP to learn to play a best response in self play (Property P1) in the 100-round IGs in which each stage game consisted of a 2x2 normal-form game conforming with each of the eight categories of games described earlier.

Figure 4.3 (left) shows the average payoff obtained by aFP in self play compared with four baseline solutions in the structured 2x2 IGs. aFP’s payoffs are always higher than the average maximin value of the stage games, but always lower than the Nash bargaining solution (NBS) [15]. Additionally, aFP’s average payoffs are typically between the average
Figure 4.3: (Left) Average payoffs obtained per round by aFP in self play in IGs of various types. aFP’s performance is compared with the average values from the stage games of (1) the maximin value, (2) the lowest NE, (3) the highest NE, and (4) the Nash bargaining solution [15] of the stage games. Results are averaged over 50 IGs each. Error bars show the standard error of the mean. (Right) The percentage of rounds that aFP played a best response in self to the action played by its partner.

Figure 4.4: Percentage of rounds in which aFP played a best response to its partner’s action in self play. Error bars (although negligible) show the standard error of the mean.

lowest-valued and highest-valued NEs. These results are driven by its aFP’s tendency to quickly learn to play a best response to its partner’s actions (Figure 4.3 right)\(^1\).

In addition to these structured IGs with 2x2 stage games, we also desired to observe how aFP’s performance scales to IGs with stage games that have more actions, as well as games in which the number of actions varied from round to round. Thus, we also evaluated this algorithm in self play in the randomly generated games of different sizes. These results are shown in Figure 4.4. When the number of actions was extended beyond two in IGs, aFP played best response less frequently compared to how often FP played the best response.
in RGs. However, these results are not surprising because (1) RGs create an unchanged environment for agents to model each other and (2) FP is not guaranteed to converge to NE.

4.4 Augmented-S#

4.4.1 Disappointment

Recall that S++ is an expert algorithm that, in each round, selects an expert to follow in the round from a set of experts $\Phi_i$. Crandall [4] showed that S++ (and hence S#) typically performs in RGs nearly as well as its best expert would have performed in the same circumstances. This concept is known as disappointment. Formally, a player’s disappointment up through round $T$ is defined as follows:

$$D_i^T = \max_{\phi \in \Phi_i} \sum_{t=1}^{T} u_i^t(\phi, \pi_{-i}^t) - \sum_{t=1}^{T} m_i(a_i^t), \quad (4.1)$$

where $\phi \in \Phi_i$ and $\pi_{-i}^t$ is a policy that agent $-i$ would have played if agent $i$ continued using expert $\phi$. $u_i^t$ represent the utility of agent $i$ at time $t$ and $m_i(a_i^t)$ is the actual payoff of agent $i$ at time $t$. The average disappointment of agent $i$ up to time $T$ is:

$$\hat{D}_i^T = \frac{D_i^T}{T} \quad (4.2)$$

As defined in earlier work by Crandall et al. [4], an agent is said to have no disappointment if

$$\lim_{T \to \infty} \hat{D}_i^T \leq 0 \quad (4.3)$$

We hold that Aug-S# should likewise demonstrate the same characteristic (Property P4). Thus, we conducted a set of simulations in IGs to determine whether this held true. Results averaged over a set of IGs with mixed payoffs (since these games consist of different

\[\text{Cyclic games have no pure-strategy NE. As such it is not possible for both players to play a best response to each others actions. This primarily accounts for the reduced percentage of best responses in the Cyclic and Mixed games.}\]
types of games) are shown in Figure 4.5. These results are encouraging as they demonstrate similar performance characteristics (in terms of disappointment) as S++ was shown to have in RGs [4]. In the majority of cases, the disappointment values decrease or converge within 1000 rounds.

4.4.2 Comparing to the Nash Bargaining Solution

In order to observe Aug-S# cooperation in self-play, we compared the average payoff values of Aug-S# with the NBS values (Property P5). Figure 4.6 shows that when cheap talk was possible, Aug-S#'s average payoffs tend to approach the NBS values of two IGs used in the user study described in the next chapter. This supported our prediction that Aug-S# cooperated well with itself in cheap talk environment. While the values are slightly lower without cheap talk (as is the case with S# in RGs) the algorithm still reaches high levels of cooperation on average in such environments. In the next chapter, we describe a user study comparing the ability of Aug-S#, aFP, and people to achieve cooperative solutions when associating with people in these same two games.
Figure 4.5: Aug-S#'s disappointment as a percentage of maximum payoff vs. different agents in mixed games.
Figure 4.6: The average payoff (as a percentage of the NBS) obtained by Aug-S# in self play in the (left) Unfair and (right) Mixed games shown in the Appendix B.
Chapter 5

User Study Design

In the previous chapter, we demonstrated via simulation that aFP and Aug-S#, our extended versions of Fictitious Play and S#, respectively, tend to maintain the same properties in many IGs as they do in RGs. Given that a primary desirable property of S# is its ability to establish cooperative relationships with people in RGs, we evaluate in this chapter whether or not Aug-S# maintains this property in IGs.

To identify the ability of Aug-S# to establish cooperative relationships with people in IGs, we designed and conducted a user study\footnote{The study was reviewed and approved by BYU IRB.} involving 72 human participants. These participants interacted with both Aug-S# and aFP in two different IGs. We describe this user study in this chapter. The results of the user study are presented in Chapter 6.

5.1 Experimental Design

The purpose of this study was to test the effectiveness of aFP and Aug-S# when developing relationships with a human partner in IGs when communication was both possible and impossible. Thus, this study had two independent variables: partner (with factor levels Human, Aug-S#, and aFP) and cheap talk (with factor levels Yes and No). When cheap talk was possible, users were allowed to send a set of pre-specified messages to each other at the beginning of each round.
<table>
<thead>
<tr>
<th>User ID</th>
<th>Opponent for Game 1</th>
<th>Opponent for Game 2</th>
<th>Cheap Talk?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>Human</td>
<td>Human</td>
<td>No</td>
</tr>
<tr>
<td>5-8</td>
<td>Aug-S#</td>
<td>Aug-S#</td>
<td>No</td>
</tr>
<tr>
<td>9-12</td>
<td>Human</td>
<td>Aug-S#</td>
<td>No</td>
</tr>
<tr>
<td>13-16</td>
<td>Aug-S#</td>
<td>Human</td>
<td>No</td>
</tr>
<tr>
<td>17-20</td>
<td>aFP</td>
<td>aFP</td>
<td>No</td>
</tr>
<tr>
<td>21-24</td>
<td>Human</td>
<td>aFP</td>
<td>No</td>
</tr>
<tr>
<td>25-28</td>
<td>aFP</td>
<td>Human</td>
<td>No</td>
</tr>
<tr>
<td>29-32</td>
<td>Aug-S#</td>
<td>aFP</td>
<td>No</td>
</tr>
<tr>
<td>33-36</td>
<td>aFP#</td>
<td>Aug-S#</td>
<td>No</td>
</tr>
<tr>
<td>37-40</td>
<td>Human</td>
<td>Human</td>
<td>Yes</td>
</tr>
<tr>
<td>41-44</td>
<td>Aug-S#</td>
<td>Aug-S#</td>
<td>Yes</td>
</tr>
<tr>
<td>45-48</td>
<td>Human</td>
<td>Aug-S#</td>
<td>Yes</td>
</tr>
<tr>
<td>49-52</td>
<td>Aug-S#</td>
<td>Human</td>
<td>Yes</td>
</tr>
<tr>
<td>53-56</td>
<td>aFP</td>
<td>aFP</td>
<td>Yes</td>
</tr>
<tr>
<td>57-60</td>
<td>Human</td>
<td>aFP</td>
<td>Yes</td>
</tr>
<tr>
<td>61-64</td>
<td>aFP</td>
<td>Human</td>
<td>Yes</td>
</tr>
<tr>
<td>65-68</td>
<td>Aug-S#</td>
<td>aFP</td>
<td>Yes</td>
</tr>
<tr>
<td>69-72</td>
<td>aFP#</td>
<td>Aug-S#</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 5.1: User Study Schedule

Each participant played two different games, and was randomly assigned two different partners. Partner orderings were counter-balanced across participants as shown for all 72 participants in Table 5.1.

For each session, participants were assigned one of the four session ID numbers. In the case of two humans playing with each other, the subject with the odd ID was the row player, while the subject with the even ID was the column player. All subjects were row players when playing against an algorithm. The condition of cheap talk or no cheap talk was also assigned to each session. There were 9 sessions for each condition. Given that each human-human interaction produces only a single data point, the study gave us 120 data points in all over the 72 subjects. When the experiment ended, each participant received an small amount of money to compensate for his or her participation. The amount was directly correlated to the payoff each participant received in the game. The payoffs in each game matrix were actually the amounts of cents a participant could receive.
As shown in Table 5.1, each session consisted of four human subjects. All participants received training prior to playing the game (which was approximately 5 to 7 minutes in duration) to make them familiar with the user interface (described in Section 5.3) and the objectives of the game. Each participant was seated such that other participants’ screens were not visible to them. Additionally, participants were requested to wear headphones throughout the study to avoid any distractions and to listen to messages when available. All participants started each game at the same time and were requested to remain quietly in their seats until every participant completed the study. After each of the two games, the participants were asked to fill out post-experimental questionnaires for that game (figure A.2 in Appendix A).

At the beginning of the user study, we also asked each of our participants to complete a demographic questionnaire (figure A.1 in Appendix A). Most participants were college students at BYU between the age of 18 to 27, with a mean age of 22.931 and a mode of 22. They came from different backgrounds and majors (Table 5.2). However, the majority of participants studied in STEM programs. 58.3% of the participants identified themselves as male while the rest identified as female.

### Table 5.2: Users’ majors information

<table>
<thead>
<tr>
<th>Majors</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics/Statistics/Computer Science</td>
<td>33</td>
</tr>
<tr>
<td>Engineering</td>
<td>12</td>
</tr>
<tr>
<td>Other STEM majors</td>
<td>2</td>
</tr>
<tr>
<td>Business</td>
<td>6</td>
</tr>
<tr>
<td>Social Sciences and Humanities</td>
<td>11</td>
</tr>
<tr>
<td>Health related majors</td>
<td>7</td>
</tr>
<tr>
<td>Design</td>
<td>1</td>
</tr>
</tbody>
</table>

5.2 The Games

As described in the previous section, each human subject played two 50-round IGs, each with a randomly selected partner (assigned according to the schedule in Table 5.1). In the first game (Game 1), each round’s payoff matrix conformed with the payoff structure of *unfair*
games as classified by Bruns [3]. In each round of this unfair game, one player had an unfair advantage over the other player with respect to payoffs as is true of all game structures in this payoff family. If one player got the best payoff, the other player would end up with the second worst payoff. This advantage changes randomly between both players over rounds.

For example, Figure 5.1 shows 2 rounds of the unfair game. The (blue) numbers with the bigger size were the payoffs of the row player while the (orange) numbers with the slightly smaller size were the payoffs of the column player. The row player had options A and B. The column player had option X and Y. In the first round, the column player had the advantage. By simply play Y, this player would receive at least the second best payoff available to himself or herself. This would force the other player to only either receive 0 or 12 which were not very desirable in respect to what was available. In this case, the row player would be forced to choose B and earned 12, the second worst payoff. The action pair B-Y was the one-shot Nash equilibrium. However, we chose this game as part of user study because it gave the players an interesting scenario.

Because the payoff matrix changed each round, it was not always in the player’s best interest for him or her to rely on the unfair advantage or the one shot Nash equilibrium. If both player chose to cooperate and played B-X in the first round and A-Y in the second round, they would be better off than playing the one shot equilibrium in both round. The row player would receive a total payoff of 55 and the column player would receive a total payoff of 20 (instead of 51 and 16, respectively).
In the second IG (Game 2), the payoff orderings of the payoff matrices for each round were randomly drawn from all types of payoff families identified by Bruns [3] which were described in the previous chapter. In short, in this game, the players encountered different situations in each round of the mixed game.

Each of these games was selected from the games generated and described in section 4.1. However, unlike those games, the chosen two games only had the first 50 rounds and all the payoff values were cut in half so that the values directly corresponded to the cents received by the players. We choose these two game because their payoff matrices vary in values across all rounds. We predicted that they would be very interesting environments for the user study. However, the study participants were not told in advance how many rounds each game would last. The complete games are shown in Appendix B.

5.3 User Interface

Study participants played the IGs in the study on desktop computers. Two different user interfaces were used depending on the communication condition assigned (cheap talk or no cheap talk). When cheap talk was allowed, the messaging function of the user interface was enabled (Figure 5.2). The participants were not given the messaging function when there was no cheap talk (Figure 5.3). In both conditions, the participants could see the game matrix and the actions available to them. They can also track the history of the game through a table which contained the previous matrices, results, and messages if applied. The interface automatically translated the payoff matrix such that all players viewed themselves as the row player.

When cheap talk was enabled, at the beginning of each round, each player could send a message to the other play by select one or more predefined speech acts. A player could also choose not to send anything by clicking the button labeled “No message.” The option to select an action was only available after both players had sent a message to the other player. After both players had made their choice, the box in the matrix that represented action pairs
5.4 Predefined Speech Acts

Participants were only permitted to communicate using a set of pre-determined phrases. This set of phrases is shown in Figure 5.4. We tried to model these messages after the previous research by Crandall et al. [5]. However, because of differences between IGs and RGs, we
had to alter several messages. We also altered and added messages after receiving feedback from people in pilot studies.

We hoped these messages would be able to help the players communicate their intentions and feelings toward one another. Some of these messages could be used to encourage one’s partner to cooperate, such as “Let’s cooperate.” Some messages allow the players to clarify the intent of the player such as “Let’s play ...” or “I should get higher payoff.” These messages also provide users with the means to express a range of emotions to help the players communicate effectively and avoid misunderstandings.

5.5 Assessing Performance

We measured the performance of our algorithms by comparing them to (i) existing results and (ii) results when humans played against each other in the same IGs. We chose the total payoffs earned over the course of the game (normalized by the NBS value of the game) as our primary performance measure.
5.6 Hypotheses

With our understanding of the algorithms given the simulation results presented in the previous chapter, along with findings in previous research [5], we made several predictions for this study. Our hypotheses were that results would be similar to the results observed in previous studies in RGs. We hypothesized that, when the other player was willing to cooperate, the payoffs received by people and Aug-S# would be close to the NBS values. Thus, when playing with another human, Aug-S# should generate payoffs that will be similar or even better than those of two human players, and should tend to approach the NBS value when cheap talk was possible. In an environment with cheap talk, the results should be better overall because it should encourage and speed up cooperation. We also predicted that, similar to self-play, Aug-S# would learn to cooperate with people quickly and would receive the highest payoffs compared to other combinations (humans and aFP).
Chapter 6

User Study Results – Interacting with People

Through the simulations reported in Chapter 4, we observed the characteristics and performance attributes of our augmented algorithms when playing against other algorithms in IGs. In this chapter, we further analyze these algorithms by observing their performance and behavior when interacting with people in the user study described in the previous chapter. We discuss the behavior of aFP and Aug-S# in these interactions with people separately, beginning with aFP. We then compare the performance of Humans, Aug-S#, and aFP in self play in the same two games used in the user study.

6.1 aFP converges to a best response when interacting with humans

Simulations reported in Chapter 4 showed that, in self play, aFP learned to play a best response to the actions played by its partner in IGs in which each game $g_i$ had a pure Nash equilibria. In this chapter, we evaluate the degree to which the algorithm also learned to play a best response when interacting with people in the user study (Property P2).

Figures 6.1 and 6.2 show the average percentage of rounds that each algorithm type played a best response to its human partner’s actions in each game with and without cheap talk. Across all conditions, aFP typically played a best response about 90% or more of the time, whereas Aug-S# and Humans played a best response to their partner’s action only about 70% of the time on average. A three-way ANOVA, with game (Unfair, Mixed Game), algorithm type (people, Aug-S#, aFP), and cheap talk (yes, no) as independent
variables, validated this trend. This analysis found a significant effect for algorithm type \( F(2, 108) = 126.81, p < 0.001 \), but not game \( F(1, 108) = 0.56, p = 0.454 \) or cheap talk \( F(1, 108) = 0.06, p = 0.803 \). Tukey pairwise comparisons between each type show that aFP played a best response more than both Aug-S# \( p = 0.001 \) and Humans \( p = 0.001 \). There was no statistical difference between Aug-S# and Humans \( p = 0.458 \). These results are shown in figure 6.3.

aFP’s best response percentage were higher than those of humans and Aug-S# in this user study. However, we wanted to find out if their rate of playing a best response when paired with people in the user study was comparable to our findings in simulation. Therefore, we proceeded to compare the best response percentages of aFP in the user study with those of aFP in self-play. Thus, we had aFP played against itself 50 times for each of the two games. The average percentages of aFP’s best response actions in these simulations and in the user study are shown in Figure 6.4. The figure shows similar rates of best response in both games under both conditions. In terms of best response, aFP performed similarly against human as it did against itself. We concluded that aFP did achieve its desired properties of converging to a best response when interacting with people.
Figure 6.2: The average percentage of rounds in which the players played a best response when paired with people in mixed game (left) with cheap talk and (right) without cheap talk. The dots represent the average best response percentages of the player types and the lines are error bars. The error bars reflect 95% confident intervals.

These percentages were good. However, we wanted to know the reason aFP did not play the best response strategy against its opponent in all rounds. When cheap talk was permitted, most of the time the humans acted according to aFPs suggestions which were typically the pure one-shot Nash equilibria of the game (when such an equilibrium existed). In some rounds (four on average per game), aFP did not communicate with its opponents because there was no pure one-shot Nash equilibria. As a result, the human player chose something different from the NE solution, such that aFP did not get the best response solutions. Without cheap talk, it was slightly harder for the humans to infer aFP’s behavior. In addition, the structure of the games contributed to aFP’s failure to achieve best response actions. In the unfair game, the game structure gave the human player a dominating action which was an advantage for the human player. However, in limited instances (four per game on average), the human did not play this dominating strategy though aFP anticipated that it would (due to human inconsistency or a lack of data about the player so far). In the mixed game, a lot of the games that aFP did not get the best response action were cyclic. These games did not have a pure one-shot NE solution. On average, there were seven rounds in the mixed games in which aFP did not play the best response action in mixed game. Out of these rounds, four of them were cyclic type (there are seven rounds drawn from this payoff
Figure 6.3: Tukey pairwise comparisons showing the average percentage of rounds in which the players played a best response when paired with people over all games and conditions. The dots represent the average best response percentages of the player types and the lines are error bars. The error bars reflect 95% confident intervals.

Figure 6.4: aFP’s best response percentage through 50 rounds in self-play (left) in unfair game and (right) in mixed game.

family in the mixed game), two games with a dominating action the human player did not play, and once in a while a human played some other unpredictable action.

These results confirm our hypothesis about aFP. The generalized algorithm did converge to a best response when playing with people despite the unpredictability of human behavior. More importantly, we observed a human tendency in IGs: they tend to aim for short term payoffs (e.g., one-shot NE), rather than more complex but more profitable cooperative actions, in more complex IGs (as opposed to RGs). This result gives insights into developing future algorithms for complex environments in the future.
6.2 Cooperation between Aug-S# and Humans in IGs

In RGs, prior work has shown that S# learns to cooperate with people and other algorithms that are inclined to cooperate [5], with cooperation rates being substantially higher when cheap talk is possible. In this section, we consider whether Aug-S# is also able to learn to cooperate with people in IGs with and without the possibility of cheap talk (Property P6). In this chapter, we do this by considering the payoffs (as a percent of the NBS value) achieved by the various player types in the user study.

The results from the user study are summarized in Figure 6.5. We ran a three-way ANOVA analysis to test the significance of our findings. Our independent variables were the (i) type of player (Human, aFP, Aug-S#), (ii) game (Unfair, Mixed), and (iii) cheap talk (yes, no). The ANOVA test was used to determine the effect of these independent variables on players’ payoff as a percentage of NBS (NBS percentage). The test showed main effects for two of these variables: type of player ($F(2, 108) = 3.45, p = 0.035$) and cheap talk ($F(1, 108) = 18.88, p < 0.001$). The interaction between these two independent variables ($F(2, 108) = 12.88, p < 0.001$) also had significant effect on the dependent variable.
Given the differences in the results with and without cheap talk, we discuss these two cases separately.

### 6.2.1 With Cheap Talk

Figure 6.5 (left) presents the differences between the algorithms’ NBS percentages with cheap talk. The plot shows that both Human and Aug-S# NBS percentages were higher than that of aFP. These values were similar between Aug-S# and Human. However, the Tukey pair-wise comparison test pointed out that the differences were not statistically significant ($p = 0.195$ for Human and aFP, $p = 0.167$ for Aug-S# and aFP). This result did not align with our expectation. This may be due to outliers, insufficient data points to discover the effect given its size, or the lack of an effect.

To better understand the results of the statistical analysis which determined the significance of the differences, we broke down the data of each interaction. Figure 6.6 gives a breakdown of each interaction in the study when cheap talk was permitted. We use the average payoffs of aFP in self-play (the myopic score) as the base-line for the comparison. If both players’ NBS percentage were a lot higher than aFP’s NBS percentages, we considered the interaction cooperative. If one player has a significantly higher NBS percentage than that of aFP while the other play’s NBS percentage value was far below that of aFP, we viewed the interaction as having resulted in the one player exploiting the other. The rest of the interactions were categorized as non-cooperative.

In the unfair game in which Aug-S# interacted with people, the majority (8 out of 12) of the interactions between Aug-S# and people were cooperative. In addition, Aug-S#’s NBS percentage in these cooperative interactions are quite high. However, the payoffs of Aug-S# in Pairings 2 and 9 (the two games in which Aug-S# was exploited by people), Aug-S#’s NBS percentages are very low. We looked into each round of these two specific interactions to learn more about these results. In Pairing 2, the human player cooperated with Aug-S# in early rounds before deciding that Aug-S# had more payoffs than he/she did.
<table>
<thead>
<tr>
<th>Pair</th>
<th>Human</th>
<th>aFP</th>
<th>NBS%</th>
<th>NBS%</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1029</td>
<td>950</td>
<td>97.54</td>
<td>95.86</td>
<td>Cooperative</td>
</tr>
<tr>
<td>2</td>
<td>863</td>
<td>628</td>
<td>81.80</td>
<td>63.37</td>
<td>Human+Exploit</td>
</tr>
<tr>
<td>3</td>
<td>955</td>
<td>923</td>
<td>90.52</td>
<td>93.14</td>
<td>Cooperative</td>
</tr>
<tr>
<td>4</td>
<td>801</td>
<td>930</td>
<td>79.52</td>
<td>93.84</td>
<td>Aug-Exploit</td>
</tr>
<tr>
<td>5</td>
<td>852</td>
<td>890</td>
<td>84.56</td>
<td>90.73</td>
<td>Cooperative</td>
</tr>
<tr>
<td>6</td>
<td>953</td>
<td>965</td>
<td>90.33</td>
<td>93.78</td>
<td>Cooperative</td>
</tr>
<tr>
<td>7</td>
<td>921</td>
<td>931</td>
<td>87.30</td>
<td>93.95</td>
<td>Cooperative</td>
</tr>
<tr>
<td>8</td>
<td>906</td>
<td>901</td>
<td>94.41</td>
<td>90.92</td>
<td>Cooperative</td>
</tr>
<tr>
<td>9</td>
<td>843</td>
<td>626</td>
<td>79.95</td>
<td>63.17</td>
<td>Human+Exploit</td>
</tr>
<tr>
<td>10</td>
<td>1058</td>
<td>966</td>
<td>100.28</td>
<td>97.48</td>
<td>Cooperative</td>
</tr>
<tr>
<td>11</td>
<td>778</td>
<td>847</td>
<td>73.74</td>
<td>85.47</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>12</td>
<td>1046</td>
<td>961</td>
<td>99.25</td>
<td>96.97</td>
<td>Cooperative</td>
</tr>
<tr>
<td>Average</td>
<td>927.92</td>
<td>877.25</td>
<td>87.9541864</td>
<td>88.5316953</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pair</th>
<th>Human</th>
<th>aFP</th>
<th>NBS%</th>
<th>NBS%</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>819</td>
<td>781</td>
<td>77.63</td>
<td>78.61</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>2</td>
<td>812</td>
<td>806</td>
<td>76.97</td>
<td>81.33</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>3</td>
<td>741</td>
<td>747</td>
<td>70.24</td>
<td>74.67</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>4</td>
<td>819</td>
<td>789</td>
<td>77.63</td>
<td>79.62</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>5</td>
<td>809</td>
<td>789</td>
<td>76.08</td>
<td>79.62</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>6</td>
<td>755</td>
<td>765</td>
<td>75.36</td>
<td>77.19</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>7</td>
<td>771</td>
<td>734</td>
<td>73.08</td>
<td>74.07</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>8</td>
<td>815</td>
<td>725</td>
<td>77.25</td>
<td>73.16</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>9</td>
<td>730</td>
<td>745</td>
<td>69.19</td>
<td>75.18</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>10</td>
<td>800</td>
<td>674</td>
<td>75.83</td>
<td>68.01</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>11</td>
<td>816</td>
<td>787</td>
<td>77.35</td>
<td>78.91</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>12</td>
<td>747</td>
<td>862</td>
<td>70.81</td>
<td>86.98</td>
<td>aFP+Exploit</td>
</tr>
<tr>
<td>Average</td>
<td>789.5</td>
<td>766.17</td>
<td>74.83</td>
<td>77.31</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pair</th>
<th>Human</th>
<th>aFP</th>
<th>NBS%</th>
<th>NBS%</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1033</td>
<td>1021</td>
<td>91.42</td>
<td>98.59</td>
<td>Cooperative</td>
</tr>
<tr>
<td>2</td>
<td>971</td>
<td>961</td>
<td>85.93</td>
<td>90.15</td>
<td>Cooperative</td>
</tr>
<tr>
<td>3</td>
<td>923</td>
<td>1059</td>
<td>81.68</td>
<td>99.34</td>
<td>Aug-Exploit</td>
</tr>
<tr>
<td>4</td>
<td>952</td>
<td>617</td>
<td>84.16</td>
<td>62.63</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>5</td>
<td>1043</td>
<td>936</td>
<td>93.30</td>
<td>87.80</td>
<td>Cooperative</td>
</tr>
<tr>
<td>6</td>
<td>1139</td>
<td>1030</td>
<td>100.80</td>
<td>96.62</td>
<td>Cooperative</td>
</tr>
<tr>
<td>7</td>
<td>928</td>
<td>867</td>
<td>82.12</td>
<td>81.33</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>8</td>
<td>1057</td>
<td>939</td>
<td>93.54</td>
<td>88.00</td>
<td>Cooperative</td>
</tr>
<tr>
<td>9</td>
<td>987</td>
<td>993</td>
<td>87.35</td>
<td>93.15</td>
<td>Cooperative</td>
</tr>
<tr>
<td>10</td>
<td>919</td>
<td>761</td>
<td>81.33</td>
<td>71.39</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>11</td>
<td>917</td>
<td>731</td>
<td>81.15</td>
<td>68.57</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>12</td>
<td>953</td>
<td>819</td>
<td>84.84</td>
<td>76.83</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>Average</td>
<td>985.08</td>
<td>900.33</td>
<td>87.18</td>
<td>84.46</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pair</th>
<th>Human</th>
<th>aFP</th>
<th>NBS%</th>
<th>NBS%</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>921</td>
<td>876</td>
<td>81.50</td>
<td>82.18</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>2</td>
<td>913</td>
<td>959</td>
<td>80.80</td>
<td>89.96</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>3</td>
<td>925</td>
<td>836</td>
<td>81.86</td>
<td>78.42</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>4</td>
<td>502</td>
<td>805</td>
<td>79.82</td>
<td>75.52</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>5</td>
<td>823</td>
<td>966</td>
<td>75.49</td>
<td>90.62</td>
<td>aFP+Exploit</td>
</tr>
<tr>
<td>6</td>
<td>858</td>
<td>794</td>
<td>75.93</td>
<td>74.48</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>7</td>
<td>864</td>
<td>942</td>
<td>76.46</td>
<td>88.37</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>8</td>
<td>852</td>
<td>891</td>
<td>75.40</td>
<td>83.58</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>9</td>
<td>905</td>
<td>928</td>
<td>80.09</td>
<td>87.05</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>10</td>
<td>936</td>
<td>848</td>
<td>82.83</td>
<td>79.55</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>11</td>
<td>883</td>
<td>1003</td>
<td>78.14</td>
<td>94.09</td>
<td>aFP+Exploit</td>
</tr>
<tr>
<td>12</td>
<td>579</td>
<td>757</td>
<td>86.64</td>
<td>71.01</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>Average</td>
<td>899.25</td>
<td>883.75</td>
<td>79.579644</td>
<td>82.903377</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.6: Payoffs for individual interactions when cheap talk was permitted.
(without actually knowing the true payoffs). The humans subsequent reaction caused Aug-S# to punish him or her. They then cooperated again and repeated the same process. Until round 42 where they just retaliate against each other for the rest of the game. In Pairing 8, Aug-S#’s punishment really hindered both players. Once again, the human player paid more attention to the ratio of giving and receiving of the rounds instead of the actual total payoffs. In this specific game, it seemed like the human player kept wanting to receive a higher payoff, which caused Aug-S# to retaliate when the solution was not what it expected. In both cases, this might not happen if the human player would have had more precise information about payoffs. Aug-S# also might have had too high of a propensity to punish. In some instances, the human players suggested to cooperate but Aug-S# continued punishing for several more rounds.

In the mixed game, Aug-S# had a harder time establishing cooperative relationships with people on average. Only half of the interactions were cooperative overall. For example, in the Pairing 4 between Aug-S# and human, the human player did not send a single message. This player completely used Aug-S#’s message to exploit the algorithm. Aug-S did punish its opponent. Hence, neither of them received good payoffs. For the non-cooperative interactions, the scenarios were the same as those in the unfair game. Both Aug-S# and the human would cooperate for several rounds, then the human player decided that he or she did not receive enough. They deviated from the solution Aug-S# proposed, leading Aug-S# to punish them. This kept repeating. In some cases, the human subjects tried to cooperate more with Aug-S# toward the end of the game. Nevertheless, it was too late to repair the damage. Despite of the challenge Aug-S# had, we could not explain why Aug-S# could not cooperate well with humans in the mixed game while all the human-human pairings achieved cooperative interactions in the same game.

Overall, across both games, aFP did not have a single cooperative interaction playing against human. On the hand, Aug-S# performed decently well when paired with many of the human players. The algorithm cooperated with its partner in the majority of cases.
In most non-cooperative interactions, Aug-S# did try to gain a better hand or showed a non-cooperative behavior through punishment. Potentially Aug-S# was either too inclined to punish its partner or it failed to properly convey that cooperation would be most beneficial to the human player.

6.2.2 Without Cheap Talk

In Figure 6.5(right), we observe that the average NBS percentage of humans and aFP were the same while the average NBS percentage of Aug-S# was substantially below the others. Tukey pairwise comparisons show that the difference between both the performance Humans and Aug-S# \((p < 0.021)\) and aFP and Aug-S# \((p < 0.001)\) are statistically significant. Thus, unlike prior results show in RGs [5], Aug-S# was substantially less successful at interacting with people in IGs than people (and aFP) are in interacting with people.

One possible explanation for this difference is that Aug-S# fails to account for changes in human behavior as the complexity of the scenario increases when they are unable to communicate with their partner. Time-varying game structures are likely perceived as more complex interactions than interactions that repeat the same scenario over and over. Given the uncertainty and greater difficulty in reasoning in these scenarios, human players potentially could become more myopic (focusing on immediate payoffs) in these situations rather than seeking long-term cooperation (as Aug-S# assumes they will).

The results shown in Figures 6.1 and 6.2 shed some insights into whether this hypothesis might have some validity. In the unfair game, humans played a best response more often without cheap talk (75% vs. 63%). However, there was very little difference in the mixed game in human best-response rates between cheap talk conditions. Overall, the difference was not statistically significant, indicating that, if this effect is indeed present, it is small enough that we did not have sufficient power in this study to observe it.

As such, we are not certain what exactly causes Aug-S# to perform so poorly when interacting with people without cheap talk in these games. Our current working hypothesis is
that people fail to understand that, in the absence of cheap talk, Aug-S#'s punishments are misunderstood, resulting in people believing the algorithm has no cohesive and cooperative strategy.

We also looked at each individual interaction across both games. Figure 6.7 shows that cooperation was hard to achieve when there was no communication in IGs for all pairings. Both human vs. human and Aug-S# vs. human pairings did much worse in term of cooperation when comparing to the interactions with cheap talk. Only one interaction between humans is categorized as cooperative. Looking at the data in Figure 6.7 for both games, we observe that Aug-S#'s NBS percentages were really low. Humans seemed to do better when playing against Aug-S# than when playing against aFP. In a substantial percentage of interactions, aFP was able to exploit people. This makes sense because communication did not affect the way aFP played. When observing the actions of Human and Aug-S# pairing, we perceived that Aug-S# still tried to choose actions that would benefit both players in the long run. On the other hand, humans played myopic actions that were most beneficial at the moment. However, because there was no communication between players, it is more difficult to understand what the players were thinking. However, it seemed like Aug-S# did followed its course and retaliate when the other player did not cooperate.

We concluded that humans were more likely to follow their partner’s cooperative strategy if the intent was clearly stated, which is done through communication. On the other hand, without knowing the opponent’s intent, humans would move toward myopic behavior, resulting in the poor performance by Aug-S# when cheap talk was not permitted.

6.3 Comparisons in Self Play

Despite Aug-S#’s struggles when interacting with humans without cheap talk, it is interesting to compare results in self play. Which of the three players (Humans, Aug-S#, or aFP) performs the best in self play in the mixed and unfair games used in the user study?
<table>
<thead>
<tr>
<th>Pair</th>
<th>Human</th>
<th>Human</th>
<th>NBS %</th>
<th>NBS %</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1121</td>
<td>792</td>
<td>206.26</td>
<td>79.92</td>
<td>Exploit</td>
</tr>
<tr>
<td>2</td>
<td>772</td>
<td>845</td>
<td>73.18</td>
<td>85.27</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>3</td>
<td>850</td>
<td>651</td>
<td>80.57</td>
<td>65.69</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>4</td>
<td>876</td>
<td>748</td>
<td>83.03</td>
<td>73.48</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>5</td>
<td>599</td>
<td>866</td>
<td>56.78</td>
<td>87.39</td>
<td>Exploit</td>
</tr>
<tr>
<td>6</td>
<td>804</td>
<td>789</td>
<td>76.21</td>
<td>79.62</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>Average</td>
<td>817</td>
<td>781.83</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pair</th>
<th>Human</th>
<th>Human</th>
<th>NBS %</th>
<th>NBS %</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>829</td>
<td>724</td>
<td>78.58</td>
<td>73.06</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>2</td>
<td>817</td>
<td>643</td>
<td>77.44</td>
<td>64.88</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>3</td>
<td>823</td>
<td>693</td>
<td>78.01</td>
<td>69.93</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>4</td>
<td>786</td>
<td>630</td>
<td>74.50</td>
<td>63.57</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>5</td>
<td>771</td>
<td>870</td>
<td>73.68</td>
<td>87.79</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>6</td>
<td>808</td>
<td>853</td>
<td>76.59</td>
<td>86.07</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>7</td>
<td>810</td>
<td>694</td>
<td>76.78</td>
<td>60.95</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>8</td>
<td>720</td>
<td>714</td>
<td>68.35</td>
<td>70.06</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>9</td>
<td>847</td>
<td>598</td>
<td>80.28</td>
<td>60.34</td>
<td>Human Exploit</td>
</tr>
<tr>
<td>10</td>
<td>771</td>
<td>712</td>
<td>73.08</td>
<td>71.85</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>11</td>
<td>793</td>
<td>806</td>
<td>75.17</td>
<td>81.33</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>12</td>
<td>817</td>
<td>714</td>
<td>77.44</td>
<td>74.07</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>Average</td>
<td>799.33</td>
<td>715.92</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pair</th>
<th>Human</th>
<th>aFP</th>
<th>NBS %</th>
<th>NBS %</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>537</td>
<td>845</td>
<td>50.90</td>
<td>85.27</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>2</td>
<td>746</td>
<td>987</td>
<td>70.71</td>
<td>99.60</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>3</td>
<td>638</td>
<td>731</td>
<td>62.37</td>
<td>73.76</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>4</td>
<td>764</td>
<td>517</td>
<td>72.42</td>
<td>92.53</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>5</td>
<td>712</td>
<td>852</td>
<td>71.38</td>
<td>85.87</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>6</td>
<td>701</td>
<td>714</td>
<td>66.45</td>
<td>72.05</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>7</td>
<td>789</td>
<td>818</td>
<td>74.79</td>
<td>82.54</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>8</td>
<td>797</td>
<td>793</td>
<td>75.55</td>
<td>80.02</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>9</td>
<td>683</td>
<td>779</td>
<td>64.74</td>
<td>78.61</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>10</td>
<td>783</td>
<td>716</td>
<td>74.22</td>
<td>76.29</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>11</td>
<td>706</td>
<td>917</td>
<td>66.92</td>
<td>94.55</td>
<td>Exploit</td>
</tr>
<tr>
<td>12</td>
<td>806</td>
<td>816</td>
<td>76.40</td>
<td>82.34</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>Average</td>
<td>726.83</td>
<td>828.75</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pair</th>
<th>Human</th>
<th>aFP</th>
<th>NBS %</th>
<th>NBS %</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>849</td>
<td>805</td>
<td>75.33</td>
<td>81.34</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>2</td>
<td>859</td>
<td>1088</td>
<td>76.02</td>
<td>100.19</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>3</td>
<td>616</td>
<td>900</td>
<td>54.51</td>
<td>84.43</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>4</td>
<td>858</td>
<td>972</td>
<td>75.93</td>
<td>91.18</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>5</td>
<td>953</td>
<td>872</td>
<td>84.34</td>
<td>76.17</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>6</td>
<td>953</td>
<td>789</td>
<td>84.34</td>
<td>70.26</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>7</td>
<td>922</td>
<td>712</td>
<td>81.59</td>
<td>66.79</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>8</td>
<td>781</td>
<td>793</td>
<td>75.55</td>
<td>80.02</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>9</td>
<td>705</td>
<td>814</td>
<td>62.39</td>
<td>85.74</td>
<td>aFP Exploit</td>
</tr>
<tr>
<td>10</td>
<td>789</td>
<td>845</td>
<td>69.82</td>
<td>79.08</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>11</td>
<td>883</td>
<td>829</td>
<td>78.14</td>
<td>77.77</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>12</td>
<td>948</td>
<td>868</td>
<td>83.89</td>
<td>81.43</td>
<td>Non-cooperative</td>
</tr>
<tr>
<td>Average</td>
<td>854.42</td>
<td>885.42</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.7: Payoffs for individual interactions when cheap talk was not permitted.
Figures 6.8 and 6.9 show the performance of each kind of player in self play. Cheap talk increased the average NBS percentage of both humans and Aug-S# in self play in these games (albeit only to a small degree in the case of humans), but not the average of aFP. Regardless, under each condition in each game, Aug-S# had the highest average, followed by humans, and then aFP. However, we note that the difference between humans and Aug-S# was not substantial in some cases.
Chapter 7

Conclusion and Future Work

While most work on learning in repeated interactions has focused on repeated games (RGs) in which the players repeatedly engage in the same scenario, such interactions do not commonly occur in reality. Thus, in this thesis, we have considered learning in repeated interactions in which the scenario can change from round to round. We have called these interactions IGs. Additionally, we have augmented two different algorithms (Fictitious Play [2] and S# [5]) previously designed for RGs so that they can be used in IGs. We then evaluated the resulting algorithms via simulation and user study in a variety of different IGs.

In simulation, our augmented algorithms maintain the same performance attributes in IGs as those of the original algorithms in RGs. The empirical distribution of aFP’s actions converge to a Nash Equilibrium and produced high percentages of best responses when playing against itself in IGs. Aug-S# tends to minimize disappointment and cooperates with a copy of itself. These results also fortified the finding in previous research on RGs [5], in which Aug-S# consistently performs well in self play simulation. In our study, it performed, on average, substantially better in self play than aFP and marginally better than humans both with and without communication.

Through the user study and the statistical analysis, we could see that both Aug-S# and aFP were trying to maintain the same performance attributes in IGs as those of the original algorithms in RGs when interacting with people. aFP did converge to a best response even with the humans’ unpredictable behavior. Its best response percentage in the user study was similar to that of aFP in self play. aFP’s best response frequencies were also
significantly higher than those of people and Aug-S#. Aug-S# did perform well against humans when cheap talk was allowed. When both humans and Aug-S# can convey their intention, the majority of the interactions between Aug-S# and humans were cooperative when communication was possible. However, when playing without cheap talk, humans appeared to have a hard time comprehending their opponents’ strategies, resulting in non-cooperative outcomes. It is especially true when there are different components to the strategies such as trigger strategies. Therefore, Aug-S# had a hard time playing against humans compared to aFP and other humans.

In summary, our augmented algorithms did carry some of the performance attributes as the original algorithms in RGs to IGs (Table 7.1). aFP had myopic behavior and converged to a Nash Equilibrium in self play. Aug-S# tends to minimize disappointment, and, in our user study, cooperated with others when they were willing to. The only exception was that when Aug-S# could not communicate its intent with its associates, it failed to produce cooperative behavior with people at levels that people achieve with each other.

This discovery has caused us to reconsider an aspect of our algorithm for future study. The trigger strategies experts in Aug-S# excessively punished their opponents when an agreement was not reached. This led to missed opportunities in several occasions. This happened in both cheap talk and no cheap talk environment. In future work, we want to avoid this excessive behavior when there was no clear agreement. We hope to find out whether humans will be able to interact with Aug-S# better when the algorithm is more lenient with punishment.
### Algorithms

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Expected Behavior</th>
<th>Our Results</th>
</tr>
</thead>
</table>
| aFP        | - The empirical distribution of actions in self play converges to a NE (P1)  
- Learn to play best response against people (P2) | - Largely confirmed for 2x2 IGs (Chapter 4.3)  
- Empirically confirmed in two 2x2 IGs (Chapter 6.1) |
| Trigger Strategies | - Offers are NE of the repeated game and are near the Pareto boundary (P3) | - Empirically confirmed across a variety of 2x2 IGs (Chapter 4.2) |
| Aug-S#     | - Obtain low disappointment when associating with many other algorithms (P4)  
- Obtain payoffs near the NBS value in self-play (P5)  
- Achieve cooperative solutions when playing with people who are willing to cooperate (P6) | - Empirically confirmed in a set of 2x2 IGs (Chapter 4.4)  
- Payoffs are 90% of NBS in typical 2x2 IGs (Chapter 4.4)  
- Mostly cooperates with cheap talk, but unsuccessful without cheap talk (Chapter 6.2) |

Table 7.1: A summary of the performance and behavioral attributes of our augmented algorithms in IGs as demonstrated by the results presented in this thesis.
Appendices
Appendix A

Questionnaire
Player ID *
Your answer

Session ID *
Your answer

How old are you? *
Your answer

What is your gender? *
- Male
- Female

In what country have you lived the longest? *
Your answer

What is your field of expertise/major? *
- Engineering
- Mathematics/Statistics/Computer Science
- Social Science
- Business
- Other: __________________________
Please describe your familiarity level with strategic games (e.g. Card games, board games, etc.)

1 | 2 | 3 | 4 | 5
---|---|---|---|---
Low | 〇 | 〇 | 〇 | 〇 | 〇 | High

Please assess yourself with respect to the following questions in the context of game-playing?

<table>
<thead>
<tr>
<th>Question</th>
<th>Never</th>
<th>Once in a while</th>
<th>Sometimes</th>
<th>Often</th>
<th>Usually</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you a cooperative person?</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Are you easily exploited?</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Are you trustworthy? (Do you live up to your words?)</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Do you use underhanded tactics to achieve your goals?</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Are you a forgiving person?</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Do you try to exploit others?</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Are you a competitive person?</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
</tbody>
</table>

Figure A.1: Demographic Questionnaire
Player ID *
Your answer

Session ID *
Your answer

Game ID *
- 0
- 1

Who do you think your game partner was? *
- A human
- A robot

How confident are you in your answer to the previous question? *

1 2 3 4 5
Just guessing

Very confident of my answer
Why did you answer question #4 like you did? (You can choose multiple answers)

- Partner’s messages
- Partner’s actions
- Responsiveness to your messages
- Time taken to cooperate with you (Long, Short)
- Other: __________

How likable was your partner? *

1 2 3 4 5

Very unlikely  O O O O O  Very likely

How well did you understand your partner’s intentions? *

1 2 3 4 5

Couldn’t understand at all  O O O O O  Understood all of them

Assuming that your partner was a human, (regardless how you answered question #4), how intelligent did your partner act? *

1 2 3 4 5

Dumb human  O O O O O  Very smart human
To what extent do the following terms describe your partner’s behavior in the game? *

<table>
<thead>
<tr>
<th>Term</th>
<th>Low</th>
<th>Medium Low</th>
<th>Medium</th>
<th>Medium High</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperative</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Devious (using underhanded tactics)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Vengeful (revengeful)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Selfish</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Forgiving</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Predictable</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Bully</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

To what extent do the following terms describe your behavior in the game? *

<table>
<thead>
<tr>
<th>Term</th>
<th>Low</th>
<th>Medium Low</th>
<th>Medium</th>
<th>Medium High</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperative</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Devious (using underhanded tactics)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Vengeful (revengeful)</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Selfish</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Forgiving</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Predictable</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Bully</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Figure A.2: Post-Experiment Questionnaire
Appendix B

The Games
Figure B.1: The Unfair Game used in the user study. Each round consisted of a normal-form game drawn from the *Unfair* payoff family [3].

<table>
<thead>
<tr>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
<th>Round 4</th>
<th>Round 5</th>
<th>Round 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>13, 2 0, 8</td>
<td>0, 7 8, 0</td>
<td>10, 0 3, 16</td>
<td>2, 10 10, 5</td>
<td>15, 3 0, 15</td>
<td>0, 24 15, 2</td>
</tr>
<tr>
<td>35, 0 12, 12</td>
<td>2, 12 30, 1</td>
<td>31, 3 0, 20</td>
<td>0, 43 1, 41</td>
<td>0, 43 1, 41</td>
<td>6, 20 13, 32</td>
</tr>
<tr>
<td>Round 7</td>
<td>Round 8</td>
<td>Round 9</td>
<td>Round 10</td>
<td>Round 11</td>
<td>Round 12</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>21, 9 15, 19</td>
<td>44, 3 33, 0</td>
<td>3, 1 1, 0</td>
<td>9, 14 0, 29</td>
<td>0, 35 5, 0</td>
<td>5, 18 4, 49</td>
</tr>
<tr>
<td>17, 14 0, 0</td>
<td>3, 9 0, 12</td>
<td>45, 9 0, 0</td>
<td>3, 9 0, 12</td>
<td>0, 0 8, 2</td>
<td>0, 0 29, 5</td>
</tr>
<tr>
<td>Round 13</td>
<td>Round 14</td>
<td>Round 15</td>
<td>Round 16</td>
<td>Round 17</td>
<td>Round 18</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>25, 25 37, 9</td>
<td>0, 35 5, 0</td>
<td>5, 18 4, 49</td>
<td>2, 27 1, 27</td>
<td>0, 15 6, 0</td>
<td>0, 15 6, 0</td>
</tr>
<tr>
<td>0, 14 50, 0</td>
<td>3, 49 18, 10</td>
<td>0, 0 8, 2</td>
<td>5, 0 16, 0</td>
<td>0, 19 12, 15</td>
<td>12, 9 3, 20</td>
</tr>
<tr>
<td>Round 19</td>
<td>Round 20</td>
<td>Round 21</td>
<td>Round 22</td>
<td>Round 23</td>
<td>Round 24</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>19, 0 0, 35</td>
<td>0, 36 33, 26</td>
<td>11, 17 0, 0</td>
<td>0, 8 34, 0</td>
<td>24, 0 5, 5</td>
<td>19, 16 2, 0</td>
</tr>
<tr>
<td>12, 23 10, 43</td>
<td>34, 0 39, 4</td>
<td>4, 36 10, 27</td>
<td>30, 32 43, 27</td>
<td>46, 1 0, 4</td>
<td>17, 31 0, 28</td>
</tr>
<tr>
<td>Round 25</td>
<td>Round 26</td>
<td>Round 27</td>
<td>Round 28</td>
<td>Round 29</td>
<td>Round 30</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>27, 4 0, 8</td>
<td>11, 0 45, 1</td>
<td>4, 3 10, 34</td>
<td>5, 9 3, 11</td>
<td>37, 3 39, 4</td>
<td>36, 2 0, 0</td>
</tr>
<tr>
<td>33, 0 25, 38</td>
<td>0, 3 33, 7</td>
<td>4, 3 10, 34</td>
<td>6, 0 4, 1</td>
<td>0, 17 15, 13</td>
<td>0, 17 15, 13</td>
</tr>
<tr>
<td>Round 31</td>
<td>Round 32</td>
<td>Round 33</td>
<td>Round 34</td>
<td>Round 35</td>
<td>Round 36</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>27, 0 36, 1</td>
<td>2, 33 1, 0</td>
<td>0, 0 26, 17</td>
<td>0, 43 1, 41</td>
<td>6, 20 13, 32</td>
<td>0, 43 1, 41</td>
</tr>
<tr>
<td>Round 37</td>
<td>Round 38</td>
<td>Round 39</td>
<td>Round 40</td>
<td>Round 41</td>
<td>Round 42</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>45, 9 0, 0</td>
<td>30, 0 13, 37</td>
<td>0, 0 30, 1</td>
<td>20, 19 11, 0</td>
<td>31, 0 0, 6</td>
<td>7, 25 15, 12</td>
</tr>
<tr>
<td>Round 43</td>
<td>Round 44</td>
<td>Round 45</td>
<td>Round 46</td>
<td>Round 47</td>
<td>Round 48</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>47, 3 12, 14</td>
<td>11, 2 8, 3</td>
<td>4, 46 22, 0</td>
<td>0, 0 21, 5</td>
<td>9, 10 1, 46</td>
<td>0, 28 20, 21</td>
</tr>
<tr>
<td>Round 49</td>
<td>Round 50</td>
<td>Round 51</td>
<td>Round 52</td>
<td>Round 53</td>
<td>Round 54</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>13, 5 15, 2</td>
<td>24, 6 43, 0</td>
<td>0, 17 15, 13</td>
<td>0, 17 15, 13</td>
<td>26, 18 47, 15</td>
<td>0, 17 15, 13</td>
</tr>
</tbody>
</table>
Figure B.2: The Mixed Game used in the user study. Each round consisted of a normal-form game randomly drawn from among the Bruns’ seven payoff families [3].
References


