Does it Pay to Know More in Games of Incomplete Information?

Olexandr Nikolaychuk

Abstract

The paper investigates experimentally the effect of information about the payoff matrix on the players' welfare attainments in repeated 2x2 games of incomplete information. Controlling for both the amount of information available to the players and the behavioral model of the opponent allows for distinct treatments of performance of informed and uninformed subjects. The main findings are as follows: (a) the informed player need not benefit from his informational advantage while the uninformed player need not be worse off; (b) asymmetric information can be used as an equilibrium selection device in games with a multiple of such; (c) asymmetric scenarios do not improve much in terms of welfare over the symmetric incomplete information baseline; (d) uninformed subjects can easily be exploited by a purposeful opponent; and in general, (e) subjects tend to perform worse than simplistic benchmark strategies.

Keywords:
learning, asymmetric information, experiments, simulations

1. Introduction

Historically, the majority of theoretical and experimental works on individual learning in games has been concerned with strategic environments of complete information where the very learning is to be understood as learning the behavioral model of the other player. One of the more recent trends in the literature is studying games of incomplete information, in which the players are deprived of the information about the payoff matrix and as such also have to learn about the game itself along the way.

This paper covers a middle ground between games of complete and incomplete information by considering scenarios with asymmetric information about the payoff matrix. I consider a number of classic 2x2 symmetric games i.e., Prisoner's Dilemma, Stag Hunt and Battle of Sexes[1] in which only one of the players knows the actual game that is being played and the other one, among other things, has to learn about it.

My primary interest lies in the effect that information about the payoff matrix can have on the welfare of the players. I therefore conduct a series of laboratory experiments in order to provide answers to the following questions:

1. Does the informed player[2] use his knowledge to his own advantage?
2. Does the presence of an informed player lead to a welfare externality on the uninformed player?
3. Does the informed player use his knowledge to increase the joint welfare[3] of the pair?
4. Do welfare gains (if any) come from lowering the degree of miscoordination or robbing the other player of payoffs?
5. How is the welfare of the uninformed player affected by the way the informed player uses his knowledge?

The experimental design includes a benchmark scenario of incomplete information about the payoff matrix and a series of asymmetric scenarios that allow for a learning environment with two principle control elements: (i) the amount of information available to the player(s) and (ii) the behavioral model of the opponent.

Some two strands of literature are related to my paper at once. Firstly, there are papers like Oechssler and Schipper (2003), and Gerber (2006), where the authors use similar setups to test how players perceive games, which they are only incompletely informed about. Secondly, there are papers like Duersch et al. (2010), Shachat and Swarthout (2002) that study how well human subjects perform against particular learning models in games of complete information. This paper can be seen as one that poses questions in the spirit of the latter for a situation inspired by the former literature.

[1]Strictly speaking, Battle of Sexes is only anti-diagonally symmetric, but that makes no difference for my particular setup as what is needed from the symmetry is that no player has a strategic advantage over the other.

[2]As opposed to the uniformed player in a symmetric game of incomplete information.

[3]While at the presence of the first two this question may at first seem redundant, it is in fact far from being so if (extra) information turns out to have opposing effects on the players' individual welfare.

In the traditional setup, each player knows his own payoff function but is unaware of that of the other one, whereas in my design, one player possesses both of these pieces of information from the start and the other one is completely ignorant. Moreover, in my learning scenario, I preclude the players from observing each other's actions. While for the completely informed player it virtually makes no difference, it makes the learning problem somewhat more difficult to tackle for his uniformed opponent. The motivation for this peculiarity is that I want to capture one of the most prominent features a learning task can have in my opinion i.e., that at any moment of time, one also has to figure out his cell in the game matrix as part of the process.

Since the behavior of the opponents needs to be controlled for, the subjects in the experiment will be playing against computer programs that represent players with different amounts of information about the payoff matrix available to them and/or level of sophistication.

In addition, I use simulations to create counterfactual welfare attainments that serve as benchmarks for the observed subject performance in each information setting.

The remainder of the paper is structured as follows. In the next section, we present the experimental design. Section 3 discusses the results. Section 4 concludes. The subject instructions are available in the appendix.

2. Experimental Design

The experimental design consists of four treatments varying in the amount of information available to the players and the learning model used for the computer opponent. In each treatment, the subjects play three blocks of 2x2 games for 50 rounds. Each block represents one of three games of interest[4], and the change of the objective game is announced to the players. The subjects are informed of the number of players and the number of actions each has in the game. They also know that their opponent is in fact a computer program following some certain algorithm in an attempt to maximize its own total payoff. The only feedback that the players receive after each period of play is their own realized payoff.

Provided that the subjects understand that it is a fixed algorithm that they are playing with, the only two possible tasks left for them (depending

[4]Subject subsamples were playing the games in all possible orders to control for possible order effects.

on the treatment) are learning the payoff matrix of the game and manipulating the inputs of the learning process on the other side. Therefore it is safe to assume that all differences in outcomes across the treatments should be attributed to how the available information about the payoff matrix is utilized by the human subjects and / or their computer opponents rather than anything else.

The baseline treatment is that of symmetric incomplete information where both human subjects and computer players do not know the objective structure of the game. The baseline computer player is programmed to follow the generic reinforcement learning technique. Another three treatments are asymmetric, in that either the human subject or computer player is endowed with information about the payoff matrix from the very beginning. In the former case, the learning algorithm of the computer player remains the same, while in the latter one it is replaced with either the basic experience-weighted-attraction (EWA) or sophisticated EWA routines[5]. For details, please consider Table 1 and Table 2.

Table 1: Experimental Design. Treatments

	Uninformed Computer	Informed Computer	
Uninformed Subject	loRE	loEWA	loEWAs
Informed Subject	hiRE	∅	∅

Treatments constituting the experiment. The first two letters in the acronyms correspond to the human subject type (i.e., *lo* and *hi* as in "low info" and "high info", respectively) and the rest of the letters correspond to the computer opponent type (i.e., *RE* and *EWA(s)* as in "reinforcement" and "experience-weighted-attraction (sophisticated)", respectively).

As it is argued in Camerer and Ho (1999), sophisticated EWA nests the original EWA and basic reinforcement learning models, which also makes the three a sensible selection from the methodological perspective. Indeed, different behavioral patterns can be achieved e.g., by restricting certain parameters of the richer model to null, which can be interpreted as dumbing down one and the same player or denying him of the knowledge about the objective structure of the game.

[5] Based on the theoretical models of Erev and Roth (1998) and Sarin and Vahid (1999, 2001), and Camerer and Ho (1999). Additional details are available in the appendix.

[6] Definitions due to Hart and and Mas-Colell (2006).

Table 2: Experimental Design. Learning Models for Computer Opponent

	Generic Name	Type	Dynamic	Learning Rule[6]	Information Used
RE	Reinforcement	Uninformed	Adaptive	Completely uncoupled	Own payoff realized
EWA	Experience-weighted-attraction	Informed	Adaptive	Uncoupled	Own payoffs, realized and counterfactual
EWAs	Sophisticated EWA	Informed	opponent model	Coupled	Own payoffs, realized and counterfactual, and opponent's payoff realized

	X	Y
X	2, 2	0, 3
Y	3, 0	1, 1

(a) Prisoner's Dilemma

	X	Y
X	2, 2	0, 1
Y	1, 0	1, 1

(b) Stag Hunt

	X	Y
X	3, 2	1, 1
Y	0, 0	2, 3

(c) Battle of Sexes

Figure 1: Games

The games of interest are the ordinal versions of Prisoner's Dilemma, Stag Hunt and Battle of Sexes as shown in Figure 1.

Stag Hunt has two Nash equilibria in pure strategies: {X, X} and {Y, Y}; and Prisoner's Dilemma has a unique Nash equilibrium in pure strategies, that is {Y, Y}. From the welfare point of view, both games are interesting in that the socially optimal upper left-hand-side cell can only be achieved (and sustained as a play outcome) if a certain amount of effort is exerted by each player. Moreover, the amount of such effort, however quantified, must be greater in the Prisoner's Dilemma than in the Stag Hunt case because strategic dominance is arguably way stronger a motive to suppress than risk dominance.

Battle of Sexes has two Nash equilibria in pure strategies: {X, X} and {Y, Y}; and either of them is socially optimal yet the players strictly but not unanimously prefer one to the other. Therefore in contrast to the other two games, there are bound to be a winner and loser in Battle of Sexes[7] and it is

[7] Unless the players coordinate on oscillating between the equilibria, which turned out to have never happened in this experiment.

a question of whether or not information can be the watershed that divides the two.

Perhaps another interesting way to look at these games is to reflect upon the nature of interaction the players are involved in in either of them. Information availability concerns aside, it is a fairly straightforward conjecture that the success of one player in Stag Hunt does not come at any cost to the other player, which is quite the opposite in case of Prisoner's Dilemma. Now in the Battle of Sexes game, it seems like the players are neither immediate antagonists nor protagonists. Whether this conjecture is true or not for symmetric incomplete information and asymmetric scenarios is an interesting empirical matter on its own and provides justification for postulating question 4 above.

From the technical perspective, the choice of games is dictated by several considerations. Firstly, there are only so many symmetric games in the 2x2 space that are economically interesting (as opposed to e.g., the game with efficient dominant strategy as defined by Eshel et al. (1998)). And secondly, I decided to exclude games that do not have Nash equilibria in pure strategies since there exists experimental evidence (see e.g., Mukherji and Runkle (2000), and Gerber (2006)) that it is generally more difficult to learn to randomize in equilibrium, and the subjects' task is already complex enough as it is.

3. Results

In total, 124 subjects took part in the experiment. At the beginning of each session, they were given some time to get to understand the instructions and ask additional questions if necessary. The instructions explained how the payoffs were determined in each round and contained a basic description of the other player. The subjects were explicitly told that the other player was a computer program trying to maximize its payoff by choosing the better action based on its past experience (and the information about the payoff matrix if available).

Before discussing the actual results, it is perhaps necessary to make sure that the human subjects indeed behave differently under various information conditions. Figure 2 provides an overview of the average empirical frequencies of observing action X across the subjects and their computer opponents per treatment per game to help understand what's going on inside each treatment and explain the results better. With loRE as the baseline, the non-parametric

Kolmogorov-Smirnov test for the equality of distributions rejects the null hypothesis at the 5% level for all treatment pairs and games but for loEWAs and loRE in Prisoner's Dilemma and loEWA and loRE in Battle of Sexes. I consider this to be a sufficient test for the differences in subject behavior, which enables me to further interpret any observed differences in welfare outcomes as those stemming from player interactions but not the noisiness factor in the opponent's learning algorithm or else.

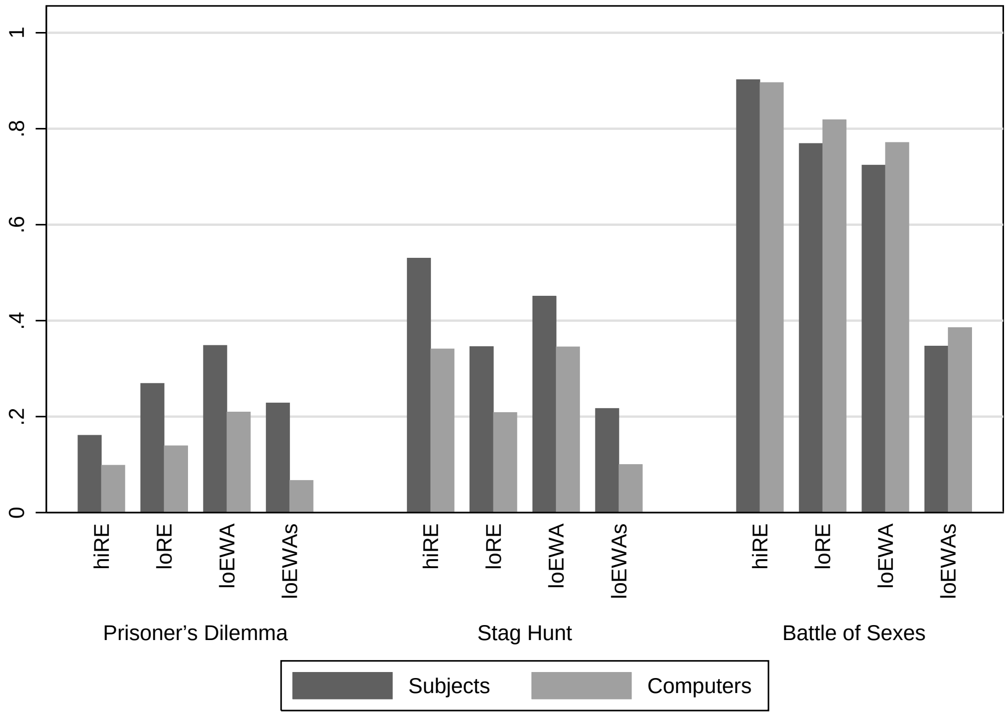


Figure 2: Average Empirical Frequencies of Action X

Means of average empirical frequencies of observing action X across the subjects and their computer opponents per treatment per game. With loRE as the baseline, the non-parametric Kolmogorov-Smirnov test for the equality of distributions rejects the null hypothesis at the 5% level for all treatment pairs and games but for loEWAs and loRE in Prisoner's Dilemma and loEWA and loRE in Battle of Sexes.

3.1. How Subjects Make Use of Extra Information

The main results of the experiment can be nicely summarized by Figure 3, where one can see the distributions of the subjects' and their opponents' average payoffs in each information treatment and game.

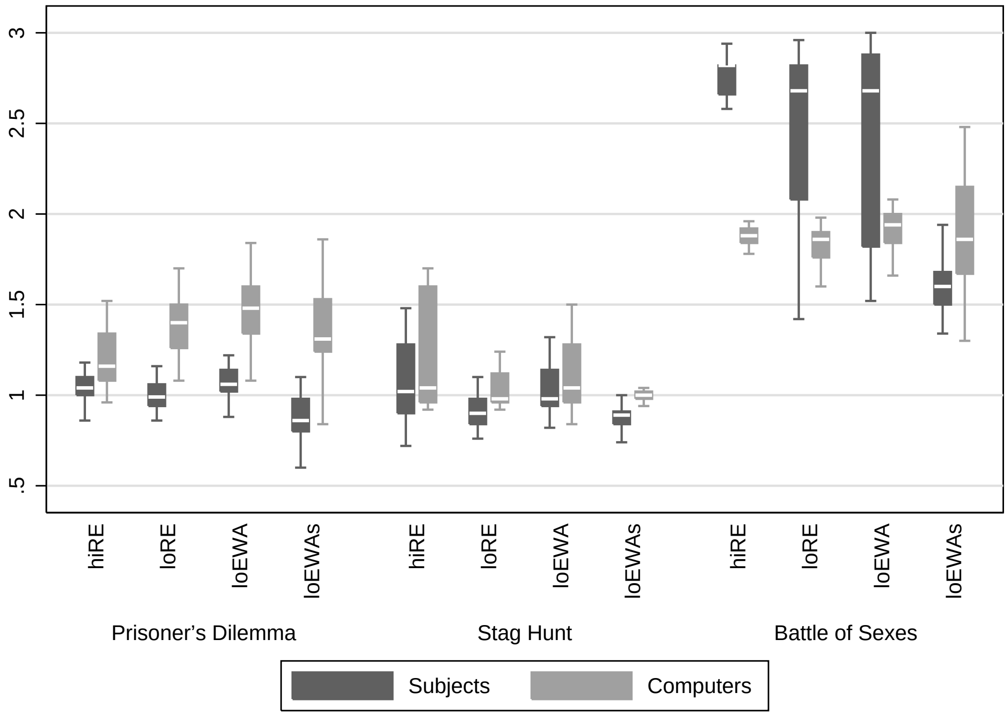


Figure 3: Average Payoff Attainments

Distribution of average contemporaneous payoffs across the subjects and their computer opponents per treatment per game. The line in the box denotes the median average profit, the boundaries of the box outline the interquartile range (IQR) between the 1st and 3rd quartiles, and the whiskers define the most extreme values within 1.5 IQR of the respective quartile.

The line in the box denotes the median average profit, the boundaries of the box outline the interquartile range (IQR) between the 1st and 3rd quartiles, and the whiskers define the most extreme values within 1.5 IQR of the respective quartile.

The accompanying Table 3 contains data on the differences in means of the subjects', opponents' and joint payoffs between the treatments in a particular pair of interest, as well as the results of the non-parametric Mann-Whitney U tests for assessing whether in one treatment the observed average payoffs tend to have larger values than in the other (double-sided p-values for the null in parentheses).

Table 3: Mean Differences in Average Payoffs

	Prisoner's Dilemma			Stag Hunt			Battle of Sexes		
	Subj.	Comp.	Joint	Subj.	Comp.	Joint	Subj.	Comp.	Joint
Δ hiRE	0.027	−0.176	−0.149	0.148	0.199	0.347	0.244	0.033	0.277
	(0.101)	(0.000)	(0.002)	(0.013)	(0.049)	(0.017)	(0.020)	(0.210)	(0.014)
Δ loEWA	0.062	0.088	0.150	0.106	0.074	0.180	−0.014	0.078	0.063
	(0.006)	(0.083)	(0.004)	(0.002)	(0.094)	(0.010)	(0.799)	(0.034)	(0.424)
Δ loEWAs	−0.104	−0.009	−0.113	−0.038	−0.058	−0.096	−0.822	0.034	−0.788
	(0.001)	(0.414)	(0.032)	(0.403)	(0.550)	(0.580)	(0.000)	(0.981)	(0.000)

Differences in means of contemporaneous payoffs across the subjects and their computer opponents per treatment per game, relative to the baseline loRE treatment. Double-sided p-values of the Mann-Whitney U test statistic corresponding to the null of no differences within a particular pair of treatments in the parentheses.

Before addressing the questions formulated in the introduction to this paper, it is interesting to note that the distributions of subjects' average payoffs tend to be dominated by those of their computer opponents regardless of the treatment in Prisoner's Dilemma and Stag Hunt. This result goes against the findings in Duersch et al. (2010) where the human subjects consistently beat their computer opponents in terms of average payoffs (and the reinforcement learning program appears to be the weakest opponent of all). In my opinion, the most crucial difference between their setup and mine is that theirs is one of complete information and mine is not. Therefore my interpretation is that simple learning algorithms can actually be quite successful[8] when the objective structure of the game is unknown. One may argue that situation in Battle of Sexes goes against this interpretation as the subjects tend to earn spectacularly more than their computer opponents in all sessions but loEWAs in this game. Yet the possible explanation would be that the classic learning models are known for their under-experimenting in early rounds of play, which in this particular case is extrememly crucial

[8]In this particular case, being successful would mean earning more than one's opponent on average.

and provides the subjects with a competitive edge. The humans must be much quicker at discovering the present equilibria and unless the computer opponent has some kind of a twist in its program as it is the case with the sophisticated version of EWA the pair becomes locked onto the equilibrium that benefits the subjects more.

The fact that the EWAs learning program has been able to beat the human subjects in what appears to be the most difficult game of the three can also be thought of as an indirect test of whether or not the basic reinforcement learning routine is a good approximation of actual human behavior in games of incomplete information. I would like to argue that the test yielded positive results, which provides an additional, ex-post justification for the selection of the opposing learning model in sessions loRE and hiRE.

Now let's get back to the original theme of the paper and compare the average payoffs in the loRE and hiRE treatments across the games.

Does the informed player use his knowledge to his own advantage? It turns out, that is not universally true. The informational leverage need not result in payoff gains as the more informed subjects in Prisoner's Dilemma do not manage to significantly increase their earnings relative to the baseline setting. They tend to defect more often and yet their opponents are not easy prey as they adjust accordingly fast enough. In the other two cases, the more informed subjects apparently use their knowledge to coordinate on the better equilibrium (from the social point of view, in case of Stag Hunt, and from the individual perspective, in case of Battle of Sexes) by committing to choosing action X more often and making the uninformed opponent follow them.

Does the presence of an informed player lead to a welfare externality on the uninformed player? Again, the answer is not unambiguous. There can be either a positive or negative externality or none for that matter. In case of Prisoner's Dilemma, the result is quite straightforward. As the more informed player tends to defect more often, there is virtually no scope for his uninformed opponent to earn more. Rather the opposite as the opponent, among other things, needs to get at least some flavor of the objective structure of the game, he is bound to lose points whenever the informed player is more inclined to defect relative to the baseline. If the actual game happens to be of the Stag Hunt type, the uninformed opponent enjoys a positive externality from the knowledge of the informed player. What is perhaps not that intuitive in this case is the fact that the uninformed player appears to benefit more than the informed one. That is, if you imagine a planner who can give knowledge about the objective structure of the game to only one of the players and prefers one of them over the other, the extra information should rather go to the less preferred one. Now in the Battle of Sexes game, even though the more informed player commits to playing action X more often thus leading the play towards the equilibrium that benefits him more, his uninformed opponent doesn't appear to be significantly affected by that in terms of payoffs in any way.

Finally, *does the informed player use his knowledge to increase the joint welfare of the pair?* As we have just seen, the answer is positive for Stag Hunt and Battle of Sexes, where the informed player is able to use extra information to his own advantage without hurting the opponent and even helping him along the way in the former case. However in case of Prisoner's Dilemma, not only wasn't the informed player successful in benefiting from his knowledge but he also managed to make his opponent significantly worse off hence decreasing the joint welfare. If the objective structure of the game happens to be that of the Prisoner's Dilemma type, it is socially optimal to destroy (asymmetric) information.

Additional light can be shed on the above findings if we look into how individual payoffs interplay in each pair as a function of information treatment. Consider Figure 4, where I present scatter plots of the subjects' average earnings against those of the computer opponents' they are matched with in a particular treatment and game. Relying on the notion of statistical association, the aim of the exercise is to estimate the extent at which one player's average payoff can be increased without having the other's average payoff decrease. If such relation were found to be statistically significant, the interpretation would be such that the players are protagonists in terms of payoffs and any payoff gains due to extra information are to be attributed to lowering the degree of miscoordination in a match. If the relation were found to be of the of the opposite sign, it would mean that the players are antagonists in terms of payoffs and are effectively hurting each other during the play.

Table 4 provides an estimate of Spearman's rank correlation coefficient, which is a non-parametric way to capture a monotonic relation, if any, between the players' payoffs in each treatment and game. The number in each cell corresponds to the estimate of coefficient; with the p-value given in the parentheses (the coefficient is not distinguishable from zero under the null). For now, we are interested in the estimates for sessions loRE and hiRE.

In the baseline treatment of symmetric incomplete information both human and computer players have to learn, which action profile yields better

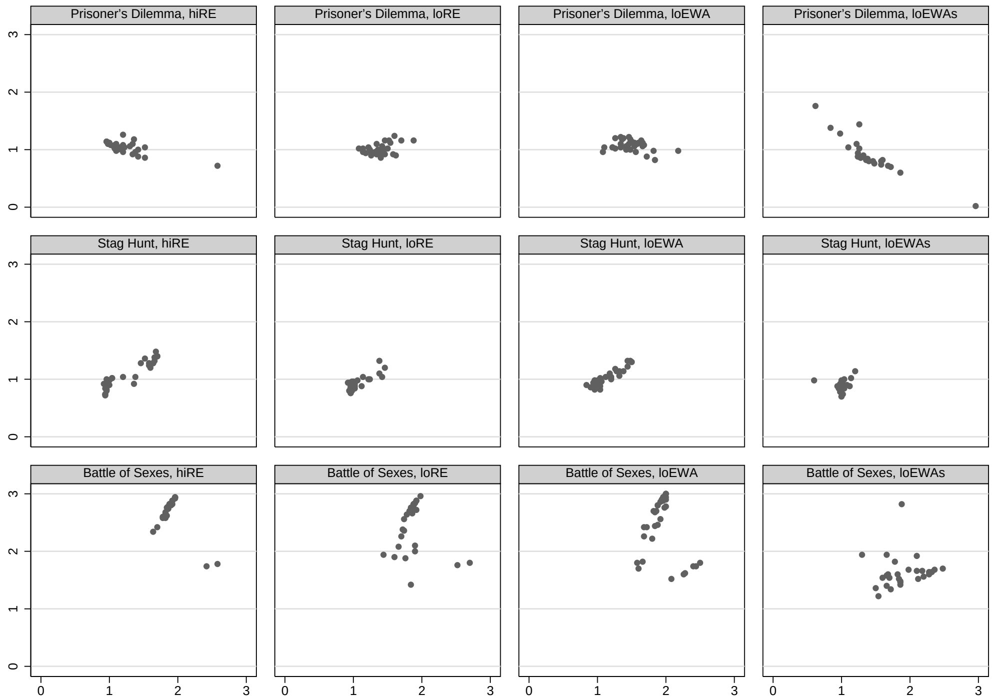


Figure 4: Average Payoffs within Subject – Computer Match

Average contemporaneous payoffs of human subjects (on the horizontal axis) against those of their computer opponent per treatment per game. Each point corresponds to a specific match, and the overall distribution indicates the degree of statistical association between the average payoffs within a pair.

payoff. It is interesting to note that in all three games the relation between individual payoffs in a match is non-negative. It is significant and positive in Stag Hunt and Battle of Sexes, which is intuitive albeit a tad bit less so in the latter case. What's even more interesting is the fact that this relation is insignificant in Prisoner's Dilemma. It appears that in games of incomplete information the learning task that the players have to solve is so complicated that on average they don't end up hurting each other in terms of welfare even when one would expect them to.

If information asymmetry is introduced to the system, the overall picture changes drastically. Although the (positive) association between the indi-

Table 4: Spearman's Rho

	Prisoner's Dilemma	Stag Hunt	Battle of Sexes
hiRE	-0.548 (0.001)	0.927 (0.000)	0.618 (0.000)
loRE	0.278 (0.137)	0.694 (0.000)	0.409 (0.025)
loEWA	-0.161 (0.370)	0.855 (0.000)	0.078 (0.665)
loEWAs	-0.939 (0.000)	0.312 (0.106)	0.330 (0.087)

Estimates of Spearman's rank correlation coefficient as a non-parametric way to capture a monotonic relation, if any, between the players' payoffs in each treatment and game. The number in each cell corresponds to the estimate of coefficient; with the p-value given in the parentheses (the coefficient is not distinguishable from zero under the null).

vidual payoffs in Stag Hunt and Battle of Sexes grows even stronger, such association becomes negative and significantly so in the Prisoner's Dilemma game. As the subjects becomes fully aware of the objective structure of the game, they start hurting their opponents in an attempt to increase own payoffs.

Now let's consider the payoff differences between the treatment pairs loEWA and loRE, and loEWAs and loRE, to see *how the welfare of the uninformed player is affected by the way the informed player uses his knowledge.*

If the uninformed subjects are playing against the EWA learner, which corresponds to an informed opponent using his knowledge in a passive or conservative way, they do not seem to be suffering any negative consequences in terms of welfare. Quite the opposite in all three games, the subjects end up being weakly better off than in the symmetric incomplete information scenario, which somewhat unexpected since there is Prisoner's Dilemma in the set of games of interest.

According to Figure 2, what happens in Prisoner's Dilemma is that the EWA computer chooses to play {X} (i.e., to cooperate) more often and induces the subjects to play {X} more often, too, and both players end up earning more on average[9] at the end of the day. Battle of Sexes is another game where the uninformed subjects can have ended up being worse off but they didn't. They were indeed induced to play {X} less often, which resulted in some payoff gains for the computer opponents, but were not significantly hurt along the way. In both games, the Spearman's rank correlation coefficient between the players' payoffs is estimated to be insignificantly different from zero, which further proves that the observed payoff gains were not at the expense of either party.

In Stag Hunt, the results are quite straightforward. The players turn out to be more successful in coordinating on the efficient outcome and both enjoy payoff gains[10] because of that. The degree of statistical association between the average payoff attainments within a pair is estimated to be highly significant and positive. Also note that, just like it was the case with treatment hiRE, the uninformed player appears to be even better off than his informed opponent on average.

The overall picture changes drastically, if the uninformed subjects are paired with the EWAs learner that uses his (asymmetric) knowledge in a sophisticated or arguably aggressive way. Relative to the benchmark scenario of symmetric incomplete information, the uninformed subjects perform weakly worse in all three games. That would have been strictly worse if not for the Stag Hunt case, where they do not appear to be affected significantly; and yet provided the nature of the game, that is a bad result in itself. The estimates of the Spearman's rank correlation coefficient are either insignificantly different from zero or different but negative (i.e., in Prisoner's Dilemma), which further implies that the uninformed player suffers a strong negative externality if (extra) information turns out to be in the wrong hands.

3.2. Welfare Benchmarks

This subsection provides additional details on how human subjects perform under various information conditions. Having learnt how human subjects perform in asymmetric treatments relative to the baseline, one may wonder to what extent they were actually successful in what they did. In a sense, can they have done better?

[9] The computer opponents are doing so only at the 10% level of significance, though.

[10] Again, the computer opponents have a hard time improving upon the 10 % significance level.

To answer that question, I construct counterfactual benchmarks by simulating average payoff attainments of several behavioral routines matched with the same learning algorithms as used in the actual experiment. The problem of choosing the most or least successful strategy goes far beyond the scope of this paper and instead, I decided to pick those that are fairly simple, fit the general learning paradigm already utilized for the computer opponents, and make some common sense altogether.

In case it is the human subjects that possess additional information about the objective structure of the game, the benchmark strategy is to always play {Y} (i.e., to defect) in Prisoner's Dilemma and {X} in Stag Hunt and Battle of Sexes. If the human subjects are the uninformed, then there are two benchmark strategies considered. The first one is the outcome of the notorious reinforcement learning routine used in the loRE treatment, and the second one is a simple randomization between the two action profiles with equal probabilities. Note that the implication of the first benchmark is that in session loRE there are two identical learning programs playing with each other, and in session loEWAs the model of the opponent assumed by the more informed computer player happens to be the correct one by construction[11]. Also note that the degenerate benchmark strategy for the informed and the second benchmark strategy for the uninformed reflect arguably the lowest bounds of performance conditional on the available information as far as one's mental effort is concerned.

Figure 5 provides the actual average payoff attainments of human subjects against the benchmark levels obtained from 10 thousand counterfactual simulations with the identical computer opponents. Benchmark A corresponds to the degenerate strategy if the subjects are the informed or to the reinforcement learning routine if they are the uninformed as explained above. Benchmark B corresponds to the equiprobable randomization between the actions for the uninformed subjects. According to the Mann-Whitney U test, the null hypothesis of the actual payoff attainments being statistically indistinguishable from either benchmark in a particular treatment and game can be rejected at the 5% confidence level for all scenarios but session loEWA in Stag Hunt (against both benchmarks).

The first thing that catches the eye here is that the subject's actual payoff

[11]Although, the loEWAs learner still cannot predict the actual realization of the probabilistic action choice subroutine of his opponent.

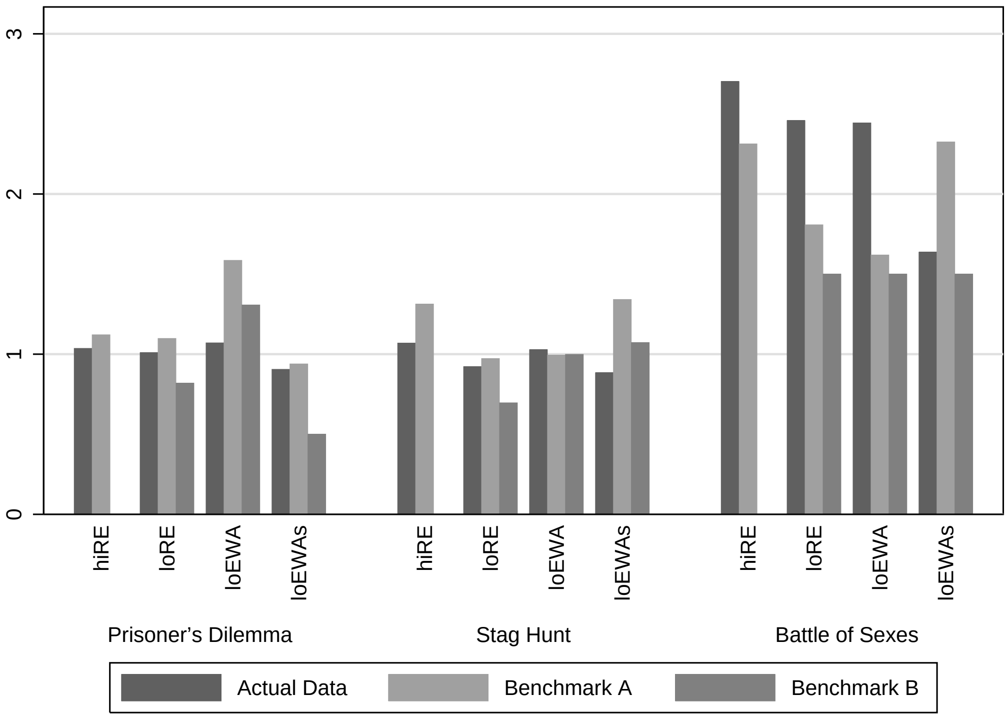


Figure 5: Average Subject Payoffs against Simulated Benchmarks

Means of average payoff attainments of human subjects against the counterfactual simulated players matched with identical computer opponents. Benchmark A corresponds to always playing {Y} in Prisoner's Dilemma, and {X} in Stag Hunt and Battle of Sexes if the subjects are the informed; or to the basic reinforcement learning routine if they are the uninformed. Benchmark B corresponds to the equiprobable randomization between the actions. The Mann-Whitney U tests reject the null hypothesis of the actual payoff attainments being statistically indistinguishable from either benchmark in a particular treatment and game at the 5% level for all scenarios but session loEWA in Stag Hunt (against both benchmarks).

attainments are again generally lower than those of the basic counterfactual behavioral routines, with the notable exception of Battle of Sexes save for the loEWAs treatment. My general explanation would be similar to that given as an interpretation of the data patterns observed in Figure 6 above. Let's take a closer look, though.

If the subjects possess (extra) information about the objective structure of the game, there is an overwhelming amount of different ways (i.e., strategies) in which that knowledge could be used. It turns out that the subjects in this experiment were successful in using their informational leverage only in case of Battle of Sexes. In Prisoner's Dilemma and Stag Hunt, they earned even less than the very simple benchmark strategies, which can be explained by certain proportions of subjects in the experiment trying to achieve the socially optimal outcome and not being successful at that in the former and not being patient enough in gearing the play towards the efficient equilibrium in the latter.

If the subjects are deprived of the information about the objective structure of the game, they perform even worse with respect to the simple benchmarks. As one carefully examines the charts going from session loRE to loEWAs, it becomes clear that as the opponent gains in knowledge, the subjects tend to lose more. There is no clear ranking as regards to the informed opponent's level of sophistication but it is obvious that the observed performance is inferior to that of the reinforcement learning routine and sometimes even to that of the overly simplistic equiprobable randomization process. Again, Battle of Sexes stands somewhat aside for the reasons discussed above.

Going back to the original question of whether the subjects can have done better or not, the answer is, certainly yes. Both the informed and uninformed subjects tend to perform poorer than fairly simple benchmark routines, and the situation appears to be significantly worse in the latter case.

4. Discussion

The general findings of my experiment can perhaps be summarized by a phrase of the sort in asymmetric scenarios when the uninformed player is deprived of knowledge of the objective structure of the game, the informed player need not benefit from his informational advantage whereas the uniformed player need not be worse off. But have we learned anything other than that?

Knowing the objective structure of the game is a very important factor in the subjects' behavior and has significant effects on the joint welfare and individual earnings within a match. Both from the social and distributional point of views, ignorance can be bliss as such asymmetric information has to be put to the proper use to yield positive results.

Playing a game of incomplete information appears to be a difficult task for the subjects as they both tend to perform inferior to some very simple learning models and, as the treatments with the informed computer opponents indicate, can be exploited easily, and that is much in contrast to the findings in the experimental literature on games of complete information. Perhaps, the only real competitive edge that the human subjects have over the traditional learning models in games of incomplete information is their non-linear propensity to experiment, which is crucial for interaction scenarios with a flavor of Battle of Sexes to them.

When the subjects are provided with the aforementioned informational advantage, but again they cannot generally beat simple benchmark routines as far as individual payoffs are concerned; so we know that they can have done better were they to behave differently. The differences in payoff outcomes between treatments loEWA and loEWAs show that the role of the informed player is an extremely important one and can have both a positive impact on an inherently antagonistic interaction scenario (i.e., Prisoner's Dilemma) or an absence of such in a protagonistic one (i.e., Stag Hunt). What we observe the informed subjects actually do is somewhere in between as they manage to make themselves better off without hurting the opponent in Stag Hunt and Battle of Sexes but fail miserably in Prisoner's Dilemma where the only thing they succeed in is producing a negative externality on the other player.

Another interesting finding is that if there is indeed scope for mutual improvement due to extra information and the more informed player commits to gearing the play towards a better outcome, then his knowledge comes at a personal cost as the uniformed opponent is likely to end up with an even greater payoff gain.

Although the objective of this study is to learn more about the value of information in general, there appear to be certain game-specific regularities I cannot but comment upon. In games with two equilibria, asymmetric information becomes a coordination vehicle that helps achieve better outcomes, either individually or socially, by lowering the degree of miscoordination between the players. However in Prisoner's Dilemma with its unique (and inefficient) equilibrium, it becomes a pure means of exploitation, and an unsuccessful one.

One last remark about the results that strikes me as peculiar is the fact that save for the Battle of Sexes game, which is arguably an easy one in terms of welfare even if at least one of the players doesn't know the payoff function, all the significant payoff differences observed are only so statistically. Yet as far as percentages are concerned, the differences are not that big. It appears that in terms of welfare, there is a huge gap between asymmetric scenarios and symmetric ones with complete information. As I argued above, having to learn the objective structure of the game must be a very difficult task for the subjects, with the manifestations of that process crowding out most of the effort by the informed player whatever it may be geared at.

5. Acknowledgements

This research was supported by grants from the CERGE-EI Foundation under a program of the Global Development Network and the Grant Agency of the Charles University in Prague. All opinions expressed are those of the author and have not been endorsed by CERGE-EI or the GDN or the Charles University.

6. References

1. Apesteguia, J., 2006. Does information matter in the commons? Experimental evidence. J. Econ. Behav. Organ. 60, 50-69.
2. Bereby-Meyer, Y., Roth, A.E., 2006. The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation. Am. Econ. Rev. 96, 1029-1042.
3. Borgers, T., Sarin, R., 1997. Learning Through Reinforcement and Replicator Dynamics. J. Econ. Theory. 77, 1-14.
4. Camerer, C., Ho, T.-H., 1999. Experience-weighted attraction learning in normal form games. Econom. 67, 827-874.
5. Chen, Y., Khoroshilov, Y., 2003. Learning under Limited Information. Games Econ. Behav. 44, 1-25.
6. Cheung, Y.-W., Friedman, D., 1997. Individual Learning in Normal Form Games: Some Laboratory Results. Games Econ. Behav. 19, 46-76.
7. Cox, J.C., Shachat, J., Walker, M., 2001. An Experiment to Evaluate Bayesian Learning of Nash Equilibrium Play. Games Econ. Behav. 34, 11-33.
8. Duersch, P., Kolb, A., Oechssler, J., Schipper, B., 2010. Rage Against the Machines: How Subjects Learn to Play Against Computers. Econ. Theory. 43:3, 407-430.

9. Erev, I., Roth, A.E., 1998. Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibrium. Am. Econ. Rev. 88, 848-881.
10. Eshel, I., Samuelson, L., Shaked, A., 1998. Altruists, Egoists, and Hooligans in a Local Interaction Model. Am. Econ. Rev. 88:1, 157-179.
11. Fehr, D., Kbler, D., Danz, D., 2008. Information and Beliefs in a Repeated Normal-form Game. SFB 649 Discuss. Pap.
12. Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Exp. Econ. 10:2, 171-178.
13. Friedman, E., Shor, M., Shenker, S., Sopher B., 2004. An experiment on learning with limited information: nonconvergence, experimentation cascades, and the advantage of being slow. Games Econ. Behav. 47, 325-352.
14. Gerber, A., 2006. Learning in and about Games. IEW Work. Pap. 234.
15. Harsanyi, J.C., 1967. Games of Incomplete Information Played by Bayesian Players, Part I. Manag. Sci. 14, 159-182.
16. Hart, S., Mas-Colell, A., 2006. Stochastic uncoupled dynamics and Nash equilibrium. Games Econ. Behav. 57, 286-303.
17. Huck, S., Sarin, R., 2004. Players With Limited Memory. Contrib. Theor. Econ. 4, Article 6.
18. Jackson, M.O., Kalai, E., 1997. Social Learning in Recurring Games. Games Econ. Behav. 21, 102-134.
19. Jordan, J.S., 1991. Bayesian Learning in Normal Form Games. Games Econ. Behav. 3, 60-81.
20. Jordan, J.S., 1993. Three Problems in Learning Mixed-Strategy Nash Equilibria. Games Econ. Behav. 5, 368-386.
21. Kagel, J.H., Kim, C., Moser, D., 1996. Fairness in Ultimatum Games with Asymmetric Information and Asymmetric Payoffs. Games Econ. Behav. 13, 100-110.
22. Kalai, E., Lehrer, E., 1993. Subjective Equilibrium in Repeated Games. Econom. 61, 1231-1240.
23. Kalai, E., Lehrer, E., 1995. Subjective Games and Equilibria. Games Econ. Behav. 8, 123-163.
24. Kaneko, M., Kline, J.K., 2008. Inductive game theory: A basic scenario. J. Math. Econ. 44, 1332-1363.

25. Mookherjee, D., Sopher, B., 1994. Learning behavior in an experimental matching pennies game. Games Econ. Behav. 7, 62-91.
26. Mukherji, A., Runkle, D.E., 2000. Learning to Be Unpredictable: An Experimental Study. Fed. Reserve Bank Minneap. Q. Rev. 24, 14-20.
27. Nicklisch, A., 2006. Perceiving Strategic Environments: An Experimental Study of Learning Under Minimal Information. MPI Collect. Goods Prepr. 17.
28. Oechssler, J., Schipper, B., 2003. Can You Guess the Game You are Playing? Games Econ. Behav. 43, 137-152.
29. Sarin, R., Vahid, F., 1999. Payoff Assessments without Probabilities: A Simple Dynamic Model of Choice. Games Econ. Behav. 28, 294-309.
30. Sarin, R., Vahid, F., 2001. Predicting How People Play Games: A Simple Dynamic Model of Choice. Games Econ. Behav. 34, 104-122.
31. Selten, R., 1975. Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. Int. J. Game Theory. 4, 25-55.
32. Selten, R., 1991. Evolution, Learning, and Economic Behavior. Games Econ. Behav. 3, 3-24.
33. Shachat, J., Swarthout, J.T., 2012. Learning about Learning in Games through Experimental Control of Strategic Interdependence. J. Econ. Dyn. Control. 36:3, 383-402.
34. Sidowski, J. B., 1957. Reward and punishment in the minimal social situation. J. Exp. Psychol. 54, 318-326.
35. Simon, H., 1957. Models of Man. Wiley, New York.
36. Slembeck, T., 1999. Low Information Games: Experimental Evidence on Learning in Ultimatum Bargaining. Univers. St. Gallen Discuss. Pap. 9903.
37. Wiseman, T., 2005. A Partial Folk Theorem for Games with Unknown Payoff Distributions. Econom. 73, 629-645.
38. Young, H.P., 2009. Learning by trial and error. Games Econ. Behav. 65, 626-643.

AppendixA. Subject Instructions

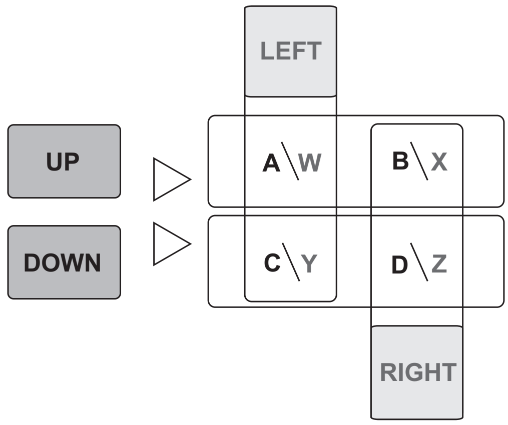
AppendixA.1. Session loRE

Dear participant,

You are about to take part in an experiment that studies how people behave in interactive environments. By choosing the right actions (to be explained later), you can significantly increase the amount of money that you will receive when the experiment is over.

The experiment consists of three stages, which we call games. We will tell you when one game is over and the next game begins.

Each game consists of 50 rounds. You will be paired with another player who will be the same throughout the whole game, and your joint decisions will determine your (and his) rewards. In each round, you will have a choice of two actions and so will the other player. Each round of the game can be represented with the help of the following picture.



In this picture, you are the red player, and the letters show rewards in each round. For example, if in some round you choose DOWN and the other player chooses LEFT, you will get C and the other player will get Y.

Neither you nor the other player will know these rewards at the beginning of the experiment. Both you and the other player will have to learn them by playing the game. We can only tell you that these rewards can be equal to 0, 1, 2 or 3 (bigger numbers mean more money) and they stay the same throughout the whole game.

During the game, you will never be told what the other player is doing. You simply make your choice, receive your reward and move to another round. The same is true for the other player. The other player will actually be a computer program that will try to earn as much money as possible throughout the experiment. The program will do its best to choose actions, which in its opinion should bring it higher rewards.

In a nutshell, you should remember the following:

- You will be playing 3 different games, one after the other;
- Each game will last for 50 rounds;
- In each round, the game remains the same;
- Your reward depends on both what you do and what the other player does;

- You will not be told what the rewards are but will have to learn them by playing the game;
- You will never be told what the other player is doing.
- The above rules also apply to the other player;
- The other player will be a computer program trying to earn as much money as possible.

AppendixA.2. Session hiRE

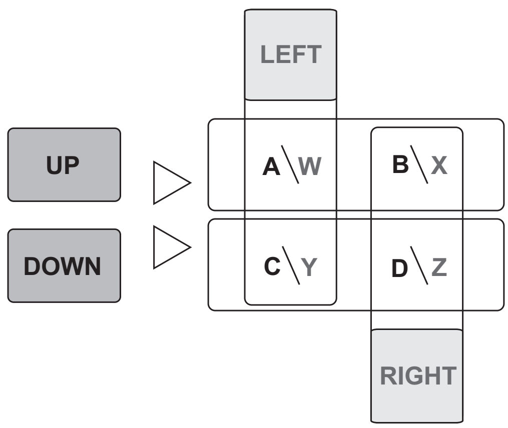
Dear participant,

You are about to take part in an experiment that studies how people behave in interactive environments. By choosing the right actions (to be explained later), you can significantly increase the amount of money that you will receive when the experiment is over.

The experiment consists of three stages, which we call games. We will tell you when one game is over and the next game begins.

Each game consists of 50 rounds. You will be paired with another player who won't be changed throughout the whole game, and your joint decisions will determine your (and his) rewards. In each round, you will have a choice of two actions and so will the other player. Each round of the game can be represented with the help of the following picture.

In this picture, you are the red player, and the letters show rewards in each round. For example, if in some round you choose DOWN and the other player chooses LEFT, you will get C and the other player will get Y.



When the game starts, you will immediately be told what the rewards are (bigger numbers mean more money), both for you and the other player. The other player will not know them from the beginning but will have to learn them by playing the game.

During the game, you will never be told what the other player is doing. You simply make your choice, receive your reward and move to another round. The same is true for the other player. The other player will actually be a computer program that will try to earn as much money as possible throughout the experiment. The program will do its best to choose actions, which in its opinion should bring it higher rewards.

In a nutshell, you should remember the following:

- You will be playing 3 different games, one after the other;
- Each game will last for 50 rounds;
- In each round, the game remains the same;
- Your reward depends on both what you do and what the other player does;
- You will never be told what the other player is doing.
- The above rules also apply to the other player;
- You will be told what the rewards are when the game starts;
- The other player will not know what the rewards are but will have to learn them by playing the game;
- The other player will be a computer program trying to earn as much money as possible.

AppendixA.3. Sessions loEWA and loEWAs

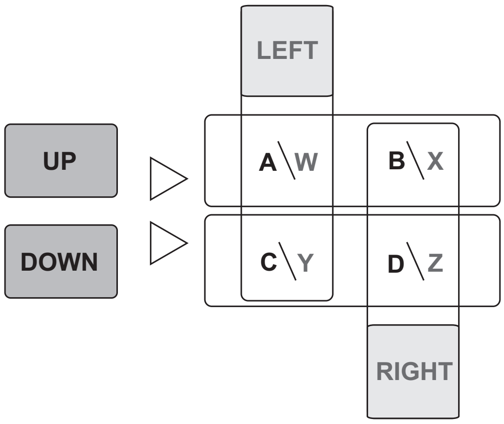
Dear participant,

You are about to take part in an experiment that studies how people behave in interactive environments. By choosing the right actions (to be explained later), you can significantly increase the amount of money that you will receive when the experiment is over.

The experiment consists of three stages, which we call games. We will tell you when one game is over and the next game begins.

Each game consists of 50 rounds. You will be paired with another player who will be the same throughout the whole game, and your joint decisions will determine your (and his) rewards. In each round, you will have a choice of two actions and so will the other player. Each round of the game can be represented with the help of the following picture.

In this picture, you are the red player, and the letters show rewards in each round. For example, if in some round you choose DOWN and the other player chooses LEFT, you will get C and the other player will get Y.



You will not know these rewards at the beginning of the game, but the other player will. You will have to learn them by playing the game. We can only tell you that these rewards can be equal to 0, 1, 2 or 3 (bigger numbers mean more money) and they stay the same throughout the whole game.

During the game, you will never be told what the other player is doing. You simply make your choice, receive your reward and move to another round. The same is true for the other player.

The other player will actually be a computer program that will try to earn as much money as possible throughout the experiment. The program will do its best to choose actions, which in its opinion should bring it higher rewards.

In a nutshell, you should remember the following:

- You will be playing 3 different games, one after another;
- Each game will last for 50 rounds;
- In each round, the game remains the same;
- Your reward depends on both what you do and what the other player does;
- You will never be told what the other player is doing.
- The above rules also apply to the other player;
- You will not be told what the rewards are but will have to learn them by playing the game;
- The other player will know what the rewards are when the game starts;
- The other player will be a computer program trying to earn as much money as possible.

AppendixB. Computer Algorithms Overview

AppendixB.1. RE Opponent

The computer opponent governed by a generic reinforcement learning routine proceeds in the following fashion:

1. Draw $\alpha_1(i), i \in \{X, Y\}$, numerical attractions to the game actions, from a uniform distribution with the support $[0, 3]$.
2. In period $n, n \in \{1, 2, ..., 49, 50\}$, choose action i according to the probability choice rule:
$$P_n(i) = \frac{\alpha_n(i)}{\alpha_n(i) + \alpha_n(-i)}.$$
3. Upon receiving payoff π_n, update the numerical attractions as:

$$\alpha_{n+1}(i) = \begin{cases} \lambda\alpha_n(i) + (1-\lambda)\pi_n & \text{if action } i \text{ has been chosen} \\ \alpha_n(i) & \text{if action } -i \text{ has been chosen} \end{cases}$$

AppendixB.2. EWA Opponent

The computer opponent governed by an experienced-weighted-attraction learning routine proceeds in the following fashion:

1. Draw $\alpha_1(i), i \in \{X, Y\}$, numerical attractions to the game actions, from a uniform distribution with the support $[0, 3]$.
2. In period $n, n \in \{1, 2, ..., 49, 50\}$, choose action i according to the probability choice rule:
$$P_n(i) = \frac{\alpha_n(i)}{\alpha_n(i) + \alpha_n(-i)}.$$
3. Upon receiving actual payoff π_n and calculating counterfactual payoff σ_n, update the numerical attractions as:

$$\alpha_{n+1}(i) = \begin{cases} \lambda\alpha_n(i) + (1-\lambda)\pi_n & \text{if action } i \text{ has been chosen} \\ \gamma\alpha_n(i) + (1-\gamma)\sigma_n & \text{if action } -i \text{ has been chosen} \end{cases}$$

AppendixB.3. EWAs Opponent

The computer opponent governed by a sophisticated version of the experienced-weighted-attraction learning routine proceeds in the following fashion:

1. Draw $\alpha_1(i), i \in \{X, Y\}$, numerical attractions to the game actions for the human player, from a uniform distribution with the support $[0, 3]$.

2. Assume that in period $n, n \in \{1, 2, ..., 49, 50\}$, the human player chooses action i according to the probability choice rule:

$$P_n(i) = \frac{\alpha_n(i)}{\alpha_n(i) + \alpha_n(-i)}.$$

3. Calculate own numerical attractions to the game actions:

$$\beta_n(i) = P_n(X)\delta_n(i, X) + P_n(Y)\delta_n(i, Y)$$

4. Choose action i according to the probability choice rule:

$$Q_n(i) = \frac{\beta_n(i)}{\beta_n(i) + \beta_n(-i)}.$$

5. Upon determining payoff π_n for the human player, update his numerical attractions as:

$$\alpha_{n+1}(i) = \begin{cases} \lambda\alpha_n(i) + (1-\lambda)\pi_n & \text{if action } i \text{ has been chosen by the human player} \\ \alpha_n(i) & \text{if action } -i \text{ has been chosen by the human player} \end{cases}$$