# Hustling in Repeated Zero-Sum Games with Imperfect Execution *

**Christopher Archibald and Yoav Shoham**
Stanford University
{cja,shoham}@stanford.edu

## Abstract

We study repeated games in which players have imperfect execution skill and one player's true skill is not common knowledge. In these settings the possibility arises of a player "hustling", or pretending to have lower execution skill than they actually have. Focusing on repeated zero-sum games, we provide a hustle-proof strategy; this strategy maximizes a player's payoff, regardless of the true skill level of the other player.

## 1 Introduction

In many games of interest agents have differing skill, and sometimes the skill of the various players is not common knowledge. Skill typically can refer broadly to any facet of a player's ability to perform both the mental and physical tasks necessary to succeed in a particular game or undertaking. One such facet, upon which we will focus, is *execution* skill.

An example of this aspect of skill can be seen in the game of billiards. A billiards player utilizes planning and reasoning skills to survey the table and decide on a shot to attempt. Execution skill determines how accurately the player is able to execute the planned shot. This type of skill clearly has an enormous impact on the success of a billiards player.

Execution skill has been studied in [Dreef *et al.*, 2002] and [Archibald *et al.*, 2010], but these works were experimental and concerned specific games. We present a model of execution skill for general games and discuss the impact that an agent's execution skill has on game situations. We then move to our main setting, repeated zero-sum games where an agent is facing an opponent with unknown execution skill. Our main result is a strategy which maximizes a player's payoff, regardless of the true execution skill level of the opponent. We then briefly discuss some related work and conclude.

## 2 Defining execution skill

One important property of execution skill, especially when execution skill is not known by opponents, is that a player may "hustle", or imitate a player with less execution skill. Conversely, of course, a less skilled player cannot imitate a player with more skill in the long run. For example, a 95% free-throw shooter in basketball could intentionally miss 45 of 100 free throws and thus appear to be a 55% shooter. But it would be highly improbable for a 55% shooter to make 95 of 100 free throws, yet this would be a likely outcome for the first player, if she were attempting to make each shot.

A natural method for reasoning about execution skill is to imagine that an agent's selected action is perturbed by the addition of noise, which results in the execution of an unintended action with some probability. A player with higher execution skill would be expected to have a higher probability of her intended action being executed, or alternatively, have her selected action be perturbed by less noise.

When a player knows the characteristics of the noise that will be applied to her action selection, then the situation is equivalent to one in which she simply selects the distribution over actions that results from the addition of noise as her original, but now mixed, strategy. Therefore, an equivalent way of modeling the execution skill of a player is to consider restrictions in the player's mixed strategy space, instead of explicitly reasoning about the addition of external noise. The player is now permitted only to select mixed strategies in the game that are "allowed", given her execution skill. Execution skill can then be represented by a specific restriction of a player's mixed strategies.

Restricting a player's mixed strategy space has been utilized as a modeling device before, both in game theory [Selten, 1975] and AI [Bowling and Veloso, 2002]. Bowling and Veloso studied limited agents in stochastic games and the player limitations, whether physical, mechanical, or mental, were modeled as restrictions in the space of mixed policies. Conditions were given which guaranteed the existence of equilibria in these limited stochastic games. Our model is a special subclass of that presented in [Bowling and Veloso, 2002], informed additionally by our intuition about the nature of restrictions that result from imperfect execution skill. The main focus of our work is execution skill in the case of incomplete information, a topic not addressed by Bowling and Veloso.

## 3 Execution skill with perfect information

We first present some preliminary game theory definitions and then return to execution skill and how it can be modeled

*This work was supported by NSF grant 0963478.

in games. Much of our notation is derived from [Shoham and Leyton-Brown, 2009].

### 3.1 Preliminaries

A *game* $G$ is a triple $(N, A, U)$ where $N$ is a finite set of $n$ players indexed by $i$, $A = A_1 \times \cdots \times A_n$, where $A_i$ is the finite set of pure strategies available to player $i$, and $U$ is a tuple $(U_1, \ldots, U_n)$, where $U_i : A \mapsto \mathbb{R}$ is player $i$'s payoff function. A subscript $-j$ denotes the set of all players except $j$.

For any set $Y$ let $\Delta(Y)$ be the set of all probability distributions over $Y$. Then the set of *mixed strategies* for player $i$ is $S_i = \Delta(A_i)$. We denote by $S = S_1 \times \cdots \times S_n$ the set of mixed strategy profiles. The *expected payoff of mixed strategy* $s$ to player $i$ ($u_i$) is calculated by $u_i(s) = \sum_{a \in A} U_i(a) \prod_{j=1}^{n} s_j(a_j)$, where $s_j(a_j)$ is the probability with which player $j$ plays action $a_j \in A_j$ under mixed strategy $s_j$.

The *safety value*, or *maxmin value* for player $i$ in game $G$ is denoted by $\underline{v}_i(G) = \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$. The safety strategy for player $i$ is simply $\alpha_i(G) = \arg\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$. The safety value is the amount that a player can guarantee himself in a game, regardless of the actions of opponents.

The *minmax value* for player $i$ in game $G$ is denoted by $\hat{v}_i(G) = min_{s_{-i}} max_{s_i} u_i(s_i, s_{-i})$. The minmax value for player $i$ is the lowest payoff that player $i$'s opponents can hold him to by any choice of $s_{-i}$, given that player $i$ correctly predicts $s_{-i}$ and plays a best response to it.

A strategy profile $s^*$ forms a *Nash equilibrium* if for every player $i$ and all strategies $s_i \in S_i$ it is the case that $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)$.

### 3.2 Modeling execution skill

As discussed earlier, imperfect execution skill can be modeled by introducing restrictions to the strategy spaces of players in games, as was used to model limitations on players in [Bowling and Veloso, 2002]. We additionally require that the restricted strategy spaces of the players be convex. This is equivalent to saying that players can utilize external randomization in selecting a strategy, and thus can always randomly select between any two allowed strategies. Once execution skill has been precisely defined, it will be clear how our model captures the afore-mentioned property of the higher skilled player being able to imitate a lower skilled player.

For simplicity of exposition and analysis, we require execution skill restrictions to be symmetric in the action space of the players. We now define execution skill in games.

**Definition 3.1.** *Given a game $G = (N, A, U)$, let $|A_i|$ be the number of actions available to player $i$ in the game. Player $i$'s* execution skill level *in $G$ is represented by $\kappa_i \in [\frac{1}{|A_i|}, 1]$, which bounds the probability that player $i$ can assign to playing any single action in $G$. This probability can be at most $\kappa_i$, and must be at least $\frac{1-\kappa_i}{|A_i|-1}$.*

Note that it follows that if player $i$ plays an action with probability $\kappa_i$, the remaining probability mass of $1-\kappa_i$ is distributed equally among the other actions. Our results can be generalized to the case where this common minimum bound does not exist, but it is included for the resulting notational simplicity. If $\kappa_i = 1$ for player $i$ in some game, then we say that player $i$ has *perfect skill* in that game. The definition also results in a total ordering over execution skill levels, as follows.

**Definition 3.2.** *If $\kappa_i > \kappa_j$, then we say that player $i$ has* more execution skill *than player $j$, while player $j$ has* less execution skill *than player $i$. If $\kappa_i = \kappa_j$, then player $i$ and player $j$ have* equal execution skill.

**Definition 3.3.** *A* game of execution skill *is a tuple $(N, A, U, K)$. $(N, A, U)$ define a game, and $K = (\kappa_1, \ldots \kappa_n)$ specifies the execution skill level for each player in the game. If $\kappa_i = 1$ for each player $i$, it is called a* game of perfect execution skill. *Otherwise, it is called a* game of imperfect execution skill.

When dealing with a game of perfect execution skill we often omit the $K$, for brevity.

**Definition 3.4.** *Given a game of execution skill $G$, we denote by $K_i$ the set of mixed strategies that player $i$ can choose in game $G$. This is the set of all mixtures that place no more than $\kappa_i$ probability on any action, and no less than $\frac{1-\kappa_i}{|A_i|-1}$.*

### 3.3 The effects of imperfect execution skill

In this section some basic analysis of a game of imperfect execution skill is presented.

#### Game transformation

When the execution skill levels of the agents are common knowledge, a game of imperfect execution skill can be rewritten as an equivalent (in expectation) game of perfect execution skill. The game of execution skill $G = (N, A, U, K)$ is transformed to $G' = (N', A', U')$, where $N' = N$, $A' = A$, and $(U_i'(a))$, is changed to

$$\sum_{y \in A} \left[ \prod_{i \in N} \left( 1\{y_i = a_i\}\kappa_i + 1\{y_i \neq a_i\} \left( \frac{1 - \kappa_i}{|A_i| - 1} \right) \right) \right] U_i(y)$$

where $1\{\cdot\}$ is an indicator function, which equals 1 when its argument is true, and is 0 otherwise.

Thus, when the players in the transformed game $G'$ play a pure joint action $a \in A'$, the resulting payoff is equivalent to what would have been obtained in expectation in the game of imperfect execution skill $G$ if each player $i$ were to play the mixed strategy which places probability $\kappa_i$ on action $a_i$, with the remaining $1 - \kappa_i$ probabilistic mass distributed evenly among the other actions.

#### Effects in zero-sum games

As prelude to the discussion of repeated zero sum games in section 4, we briefly discuss the impact that a player's execution skill level has on the value of a zero-sum game.

The famous minimax theorem [von Neumann, 1928] states that in any finite, two-player, zero-sum game $G$, in any Nash equilibrium each player receives a payoff $v_i(G)$ that is equal to both his maxmin value and his minmax value ($v_i(G) = \hat{v}_i(G) = \underline{v}_i(G)$. For this reason we focus on the impact that execution skill has on a player's safety value in a game. Due to its fairly straightforward nature, we omit the proof of the next theorem.

**Theorem 3.5.** *Let $G'$ and $G''$ be games of imperfect execution skill which differ only in the execution skill of player $i$. Assume that $\kappa_i' > \kappa_i''$. Then the safety value of player $i$ in $G'$, $\underline{v}_i(G')$, is greater than or equal to the safety value of player $i$ in $G''$ ($\underline{v}_i(G') \geq \underline{v}_i(G'')$).*

Thus, greater execution skill for a player can never decrease their safety value in a particular game. This fact gives intuition for the setting described in the next section. Many other theorems can be stated regarding the effect of imperfect execution skill in one-shot normal form games, but we omit these due to space limitations.

# 4 Repeated games of incomplete information

We now come to our main focus, repeated game settings where execution skill is not common knowledge. As a motivating example, consider the setting presented in the 1961 film "The Hustler". In the film, Jackie Gleason's character, Minnesota Fats, is a famous champion of billiards. He has participated in many publicly observed billiards games and everyone is aware of his very high execution skill level. The film introduces Eddie Felson, the character played by Paul Newman, as a challenger to Minnesota Fats who desires to beat him in a head-to-head match. In this case the execution skill of Eddie Felson is not known to Minnesota Fats, whereas Eddie has knowledge of Minnesota's skill level. As the name of the movie indicates, one common strategy for unknown players in such situations is to hustle, pretending to have less execution skill than they actually have. The aim of such behavior is to get the opponent to utilize a strategy that he normally wouldn't, which can lead to the opponent receiving a lower payoff than he should in the game.

A strategy for an agent facing an opponent of unknown execution skill in a game has two seeming opposing desiderata. First, the strategy should not be susceptible to hustling, where the agent loses payoff due to the opponent's initial imitations of low skill. Second, the strategy should not overestimate the opponent's execution skill and consequently cause the player to fail to receive his maximal possible payoff given his opponent's actual skill level. In many settings maximizing the payoff obtained against any opponent is crucial to overall success.

The specific setting we investigate is two-player infinitely repeated zero-sum games of imperfect execution skill. In this setting player 2 retains the same execution skill level throughout the infinite duration of the repeated game, and player 1 has no knowledge of this level. Conversely, it is common knowledge between the players that player 1 has perfect execution skill. The players cannot communicate outside of the game, but can observe the actions taken by each player in all previous rounds. This repeated setting gives player 1 a chance, throughout the course of the game, to learn about the execution skill level of player 2.

## 4.1 Preliminaries

For each $t = 1, 2, \ldots$, let $H^t$ be the set of possible histories up to (but not including) stage $t$, namely $H^t = (A \times \ldots \times A(t\text{-}1 \text{ times}))$. The empty history will be designated by $h^0$. A pure strategy for player $i$ is a collection $\sigma_i = \{\sigma_i^t\}_{t=1}^{\infty}$, where $\sigma_i^t \colon H^t \mapsto A_i$. A mixed strategy for each player is a probability distribution over pure strategies. Since our setting assumes perfect recall, we can focus our attention on behavioral strategies, as shown in [Kuhn, 1953]. Behavioral strategies are strategies where the players randomize independently at each stage of the game. The players' behavioral strategies can be defined as $\sigma_1^t \colon H^t \mapsto \Delta(A_1)$ and $\sigma_2^t \colon H^t \mapsto K_2$. This ensures that player 2 selects only valid mixtures given his execution skill level.

We now define the payoffs in the infinite repeated game. For a history $h^T = ((a_1^1, a_2^1), (a_1^2, a_2^2), \ldots (a_1^{T-1}, a_2^{T-1}))$, we define the payoff to each player $i$ as

$$\beta_i^T = (\frac{1}{T}) \sum_{t=1}^{T} u_i(a_1^t, a_2^t)$$

A pair of strategies $\sigma = (\sigma_1, \sigma_2)$ by the two players induces a probability distribution over histories. Denote by $E_{\sigma,p}(\beta_i^T)$ the expected value of $\beta_i^T$ with respect to this induced distribution.

A pair of strategies $\sigma = (\sigma_1, \sigma_2)$ form a Nash equilibrium if for each player $i$ it is the case that for all strategies $\hat{\sigma_i}$ of player $i$

$$\lim_{T\to\infty} \sup E_{\sigma,p}(\beta_i^T) \geq \lim_{T\to\infty} \sup E_{(\hat{\sigma}_i,\sigma_{-i}),p}(\beta_i^T)$$

## 4.2 Zero-sum repeated games of complete information

We first briefly discuss what the value of the infinite repeated game would be if player 2's execution skill level were common knowledge. We let $G_x$ refer to a single stage zero-sum game of imperfect execution skill where player 2 has an execution skill level of $x \in [\frac{1}{|A_2|}, 1]$ and player 1 has perfect execution skill, and where $x$ is common knowledge. The infinite repetition of $G_x$ is denoted by $G_x^\infty$. We will denote by $\mathcal{G}_x^\infty$ the repetition of $G_x$ when $x$ is only known to player 2. Recall from the discussion in section 3.3 that there is a well-defined value for each player in $G_x$,

$$v_i(G_x) = \underline{v}_i(G_x) = \hat{v}_i(G_x).$$

**Theorem 4.1.** *Given common knowledge of $x$, the value of the infinite repeated game $G_x^\infty$, is $v_i(G_x)$ for each player $i$.*

*Proof.* From the folk theorem for repeated games [Fudenberg and Tirole, 1991] we know that in any Nash equilibrium of a repeated game, each player will receive a payoff of at least $\hat{v}_i(G_x)$. The only value which achieves this for each player, since the game is zero-sum, is the value of the stage game $v_i(G_x)$. This value can be obtained by playing an equilibrium strategy of the stage game $G_x$ in each round of the repeated game. □

## 4.3 From complete to incomplete information

Our goal is now to discover strategies for each player which form an equilibrium in the repeated game $\mathcal{G}_x^\infty$, and guarantee that each will receive the $v_i(G_x)$ in the infinitely repeated game, as if $x$ were commonly known. Just attaining $v_2(G_x)$ is simple for player 2, as he knows the value of $x$ and can simply play the safety strategy of the game $G_x$, ensuring that he gets at least $v_2(G_x)$, regardless of the actions of player 1. Two issues remain. What strategy if any for player 1 also obtains this value, and will these two strategies form an equilibrium?

To simplify the following notation, in what follows we assume that $|A_i| = 2$, although all our results extend simply to the general case. As $x$ is varied between 0.5 and 1, player 1 can have at most two different safety strategies, $\alpha_1(G_{0.5})$ and $\alpha_1(G_1)$. Let $q^*$ be the value of $x$ at which these strategies both yield the same safety value. If $x < q^*$, then the safety strategy for player 1 is $\alpha_1(G_{0.5})$, while if $x > q^*$ it is $\alpha_1(G_1)$. If player 1 has a single safety strategy for all possible values of $x$ (i.e. $\alpha_1(G_{0.5}) = \alpha_1(G_1)$), then she can simply play this strategy and always be guaranteed $v_1(G_x)$. With that trivial solution discussed, we assume in what follows that $\alpha_1(G_{0.5}) \neq \alpha_1(G_1)$.

### An abstract analysis

At a high level there are two different approaches that player 1 could use. Player 1 can begin the game either by playing $\alpha_1(G_{0.5})$ or $\alpha_1(G_1)$ and then modify her strategy choice in future rounds depending on player 2's actions. Recall from theorem 3.5 that $\underline{v}_2(G_x)$ will be at its minimum value (and $\underline{v}_1(G_x)$ at its maximum value) when $x = 0.5$, and $\underline{v}_2(G_x)$ will be at its maximum value (and ($\underline{v}_1(G_x)$ at its minimum value) when $x = 1$. Also, when playing $\alpha_1(G_1)$ player 1 is guaranteed exactly $v_1(G_1)$ in expectation, regardless of the actions of player 2.

It is clear then, that if player 1 were to begin by playing $\alpha_1(G_1)$, player 2 could respond by randomizing equally between his two actions, regardless of what his true execution skill level were. This would make it impossible for player 1 to tell from the action choices of player 2 anything about the value of $x$, leaving player 1 nothing to base a change of strategy upon. It seems clear that this approach by player 1 would lead to an average payoff for player 1 in the repeated game of $v_1(G_1)$, which is the best payoff possible for player 2.

However, if player 1 begins by playing $\alpha_1(G_{0.5})$ then player 2 is forced, in a certain sense, to reveal his execution skill level by playing his best response often enough to get player 1 to change her strategy choice. Let $\gamma_2$ be player 2's pure strategy best response to strategy $\alpha_1(G_{0.5})$, and $\hat{\gamma}_2$ player 2's other pure strategy. Throughout the duration of the game, player 1 can keep track of $r_n$, the number of times that player 2 plays action $\gamma_2$ in the first $n$ rounds. Thus, $r_n$ will always be less than or equal to $n$.

If, in response to player 1's strategy, player 2 never plays $\gamma_2$ more frequently than $q^*$ ($\frac{r_n}{n} < q^* \forall n$), then player 1's payoff in the repeated game will be greater than $v_1(G_1)$. Of course, player 2 can only reliably play $\gamma_2$ more frequently than $q^*$ if $x > q^*$. When $\frac{r_n}{n}$ surpasses $q^*$, player 1 can modify her strategy. The difficulty, as we shall see, lies in switching between the two strategies in such a way that does not leave player 1 vulnerable to exploitation.

### Example of a naive solution failing

Let's see what can happen if player 1 uses a naive strategy based on the previous idea, playing $\alpha_1(G_{0.5})$ in round $n$ if $\frac{r_n}{n} \leq q^*$ and $\alpha_1(G_1)$ otherwise. We consider the following game, where $v_1(G_1) = 2.5$, $\alpha_1(G_{0.5}) = D$, $\alpha_1(G_1) = (U : 0.5, D : 0.5)$, $q^* = 0.75$ and $\gamma_2 = R$.

| | L | R |
|---|---|---|
| U | 1,-1 | 3,-3 |
| D | 4,-4 | 2,-2 |

Figure 1: Example zero-sum game

When player 2 has perfect execution skill ($x = 1$) there is a sequence of actions ($a_2^n$) which yield player 1 a lower payoff than $v_1(G_1) = 2.5$. We show this sequence, along with, for each round, player 1's strategy choice $s_1^n$, $\frac{r_n}{n}$, and the expected payoff to player 1 in that round ($EU_1^n$).

| $n$ | $s_1^n$ | $a_2^n$ | $\frac{r_n}{n}$ | $EU_1^n$ |
|---|---|---|---|---|
| 0 | D | R | - | 2 |
| 1 | $\alpha_1(G_1)$ | L | 1.0 | 2.5 |
| 2 | D | R | 0.5 | 2 |
| 3 | D | R | 0.66 | 2 |
| 4 | D | R | 0.75 | 2 |
| 5 | $\alpha_1(G_1)$ | L | 0.8 | 2.5 |
| 6 | D | R | 0.66 | 2 |
| 7 | D | R | 0.714 | 2 |
| 8 | D | R | 0.75 | 2 |
| 9 | $\alpha_1(G_1)$ | L | 0.77 | 2.5 |

The cyclical nature of player 2's exploitation of player 1's strategy is immediately apparent. This exploitation enables player 2 to get a higher payoff than he should in the game. This example shows that player 1's strategy must be carefully constructed to ensure that such exploitation does not occur.

### A provably correct solution

We represent player 1's strategy by a sequence of functions $\{f_n\}$, where each $f_n : \mathbb{N} \mapsto [0, 1)$. Each function $f_n$ will be a mapping from valid values of $r_n$ to values in $[0, 1)$. At any point in the game, player 1 will play according to the function $f_n$, where $n$ is the current round number. In each round player 1 will play the strategy $\alpha_1(G_1)$ with probability $f_n(r_n)$ and strategy $\alpha_1(G_{0.5})$ with probability $1 - f_n(r_n)$.

**Definition 4.2.** *We call a sequence of functions* $\{f_n\}$ Hustle-proof, *if*

1. $f_0 = 0$
2. $k < d \implies f_n(k) \leq f_n(d)$
3. *If* $r_n \leq q^* \cdot n$, *then* $f_n(r_n) = 0$
4. $f_n(nq^* + k) \downarrow 0$ *as* $n \to \infty$, *for all* $k \in \mathbb{N}$
5. $f_n(k \cdot n) \uparrow 1$ *as* $n \to \infty$, *for all* $k \in (q^*, 1]$
6. *If* $\{k_n\} \to q^*$, *then* $\frac{f_{n+1}(k_n n+1) - f_n(k_n n)}{f_n(k_n n) - f_{n+1}(k_n n)} \to \frac{1-q^*}{q^*}$.

Before we move to our main result, we give an example of a hustle-proof sequence of functions.

$$w_n(r_n) = \begin{cases} 0 & r_n \leq nq^* \\ 1 - \frac{1}{1+\frac{r_n - nq^*}{\sqrt{n}}} & r_n > nq^* \end{cases} \tag{1}$$

**Theorem 4.3.** *If player 1's strategy corresponds to a hustle-proof function sequence she is guaranteed at least $v_1(G_x)$ in the repeated game $\mathcal{G}_x^\infty$.*

*Proof.* We proceed to show that for any sequence of actions by player 2, the value of the repeated game where player 1 uses the proposed strategy will be at least $v_1(G_x)$. Since player 1's strategy depends upon $n$ and $r_n$, we characterize player 2's action sequences by the behavior of the ratio $\frac{r_n}{n}$.

[**Case 1**] There exists a $T$ and a $k \leq q^*$ such that $\forall t > T$, $\frac{r_t}{t} \leq k$.

By property 3 we also then have that player 1 will be playing $\alpha_1(G_{0.5})$ after the point $T$. Player 1's average payoff will converge to at least $v_1(G_k)$. For any skill level $x > k$ of player 2, this ensures player 2 at least $v_1(G_x)$ as desired. Notice that if $x < q^*$, then by the definition of execution skill and the law of large numbers, the observed frequency with which player 2 plays action $\gamma_2$ can converge to at most $x$. Thus, any sequence created by such a player must fall into this case, guaranteeing player 1 at least $v_1(G_x)$.

[**Case 2**] There exists a $T$ and a $k > q^*$ such that $\forall t > T$, $\frac{r_t}{t} \geq k$.

Since for $n > T$ we know that $\frac{r_n}{n} \geq k > q^*$ holds, by properties 2 and 5 we know that $f_n(r_n) \to 1$, which means that player 2 will play the strategy $\alpha_1(G_1)$ with probability converging to 1. This strategy guarantees player 1 a payoff of $v_1(G_1)$, which is equal to $v_1(G_x)$, since player 2 can only create such a sequence when $x > k$.

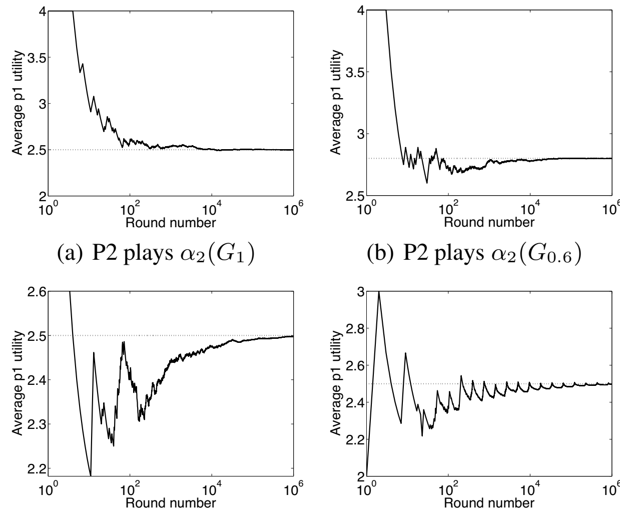
[**Case 3**] For any $T$ and any $k > q^*$, there exists a $t > T$ such that $\frac{r_t}{t} < k$.

The only types of action sequences not handled by the previous cases are those which will cause the ratio $\frac{r_n}{n}$ to return arbitrarily close to $q^*$ an infinite number of times. We note that for this to occur it must be the case that $x \geq q^*$. In each of these cases $v_1(G_x) = v_1(G_1) = v_1(G_{q^*})$.

In this case we need to ensure that the magnitude of any gains by player 2 will decrease to 0 over time. In a round where player 2 plays $\gamma_2$, player 1 has expected payoff of less than $v_1(G_x)$, while if player 2 plays $\hat{\gamma}_2$, player 1 receives more than $v_1(G_x)$. For any $k > q^*$, if there is an infinite subsequence of $\frac{r_n}{n}$ which is above $k$, then by case 2 this subsequence will eventually yield an average payoff of $v_1(G_1)$, as desired. The other subsequence of $\frac{r_n}{n}$, which must converge to $q^*$, will by properties 4 and 6 converge to yielding an average reward of $v_1(G_1)$.

Thus, we have shown that for any action sequence and possible skill level $x$ of player 2 given that action sequence, the proposed strategy guarantees player 1 an expected payoff in the infinite repeated game of $v_1(G_x)$. □

Figure 2 shows some example payoff sequences corresponding to the example game in figure 4.3, when player 1 utilizes the strategy corresponding to equation 1. Each graph demonstrates the convergence of the average payoff of strategy $\{w_n\}$ against different opponent strategies. In figure 2(a), player 2 plays a constant strategy of $\alpha_2(G_1)$ in each round, and the average payoff converges to $v_1(G_1) = v_1(G_{q^*}) = 2.5$. In figure 2(b), player 2 plays the constant strategy of $\alpha_2(G_{0.6})$ in each round, and the average payoff converges to $v_1(G_{0.6}) = 2.8$. In figure 2(c), player 2 plays $\gamma_2$ with probability 0.85 in each round, and the average payoff again converges to $v_1(G_1) = 2.5$. In the final figure, 2(d), player 2 utilizes a strategy which causes $\frac{r_n}{n}$ to cycle higher than $q^* = 0.75$ and then back down to $q^*$. In each graph the value to which the average payoff should converge is shown by a dashed line.




(a) P2 plays $\alpha_2(G_1)$ (b) P2 plays $\alpha_2(G_{0.6})$

(c) P2 plays $\gamma_2$ w.p. 0.8 (d) P2 plays cyclically

Figure 2: Average P1 payoff against different P2 strategies

**Corollary 4.4.** *The following two strategies are in equilibrium in the game $\mathcal{G}_X^\infty$:*

- *player 1 uses a strategy corresponding to a hustle-proof sequence of functions*
- *player 2 plays $\alpha_2(G_x)$, his safety strategy in the game where $x$ is common knowledge*

*Proof.* Follows directly from the fact that the strategy $\alpha_2(G_x)$ by definition guarantees player 2 at least $v_2(G_x) = -v_1(G_x)$, while player 1's strategy guarantees her at least $v_1(G_x)$, by theorem 4.3. □

One thing we can consider is how much information player 1 learns throughout the entire course of the game about player 2's execution skill level. If $x < q^*$, then player 1 will learn exactly what the execution skill level of player 2 is, because the limit average payoff $v_1(G_x)$ achieved in equilibrium at each such skill level is unique. If $x > q^*$, player 1 will not know anything more than this fact. Since player 2's strategy is the same at each of these skill levels, the only information player 1 gets is that player 2 has enough execution skill to secure his safety value in the game $G_1$.

At this point we can say a few words about how this method extends to games beyond $2 \times 2$. It these games, instead of a single $q^*$, there will be a finite set of $q^*$ values, since there can be multiple execution skill values for player 2 where player 1's safety strategy changes. These $q^*$ values will divide up player 2's execution skill range into intervals, where player 1's safety strategy is constant on each interval. Player 1 will still begin by playing her safety strategy assuming that player 2 has minimal execution skill, forcing player 2 to reveal the fact that he in fact has more execution skill. At any point in time player 1 will use a hustleproof function to randomize between the safety strategy corresponding to the interval in which the current skill estimate ($\frac{r_n}{n}$ in this paper) lies and the safety strategy of the next lower interval.

## 5 Related work

Work has been done in the game theory community where players are considered to "tremble". One example is the work regarding trembling-hand perfect equilibria [Selten, 1975]. However, the mathematics involved only treats these errors in the limit as their magnitudes approach zero. Work that investigates finite trembles, or execution error, can be found in that on proper equilibria [Myerson, 1978], quantal response equilibria [McKelvey and Palfrey, 1995] and imperfect equilibria [Beja, 1992]. The goal of these works is to refine notions of equilibrium. In each of these works the trembles of the players are correlated with the expected payoff each of their actions will receive, and so are only clearly defined given a strategy profile for the opposing players. We view imperfect execution skill in a more fundamental way, inherently related only to the actions possible and the action being attempted, not with the payoff each action might garner. For example we wouldn't expect a reward of one million dollars for the 55% shooter in the basketball example to have an impact on the probability that he could make 95 of the 100 shots.

Other work has focused on repeated games of incomplete information. The first work in the area can be found in [Aumann and Maschler, 1995]. In these games there is a finite set of $k$ normal form games, one of which is chosen according to a distribution $p$, and that game is played repeatedly. Much of the subsequent work in the area (e.g. [Hart, 1985; Shalev, 1994]) focuses, like us, on the setting where the players have different information about the actual game being played. In that setting, one player is informed of the game being played, while the other is not. Recent work within the multiagent community [Gilpin and Sandholm, 2008] addressed these games with the purpose of calculating an equilibrium strategy in the repeated game for the players.

## 6 Conclusion

In this paper we introduced the concept of execution skill and presented a hustle-proof strategy for an agent in a repeated zero-sum game of imperfect execution skill, when that agent does not know the execution skill level of her opponent. This strategy is robust in the sense that it ensures that the agent achieves the same value that she would achieve in the game if she knew exactly the opponent's skill level. This strategy in essence forces the opponent to exhibit his execution skill in the game since otherwise his payoff will suffer. This high level approach in zero-sum games with imperfectly skilled opponents can give insight into other situations where participants in multiagent settings have imperfect execution skill.

In the future we intend to investigate similar strategies for non-zero-sum games, enabling agents to avoid being hustled, and indeed to gain when it is advantageous. We hope this work lays a foundation for future investigations into other aspects of execution skill, including incorporating models of changing execution skill in repeated game settings. We can also evaluate payoffs differently in the infinite game, using discounted future rewards, and we plan to compare successful strategies in that setting with the results shown here.

## References

[Archibald *et al.*, 2010] Christopher Archibald, Alon Altman, and Yoav Shoham. Success, strategy and skill: an experimental study. In *Proceedings of AAMAS 2010*, 2010.

[Aumann and Maschler, 1995] R. Aumann and M. Maschler. *Repeated Games with Incomplete Information*. The MIT Press, 1995.

[Beja, 1992] Avraham Beja. Imperfect equilibrium. *Games and Economic Behavior*, 4(1):18–36, 1992.

[Bowling and Veloso, 2002] Michael Bowling and Manuela Veloso. Existence of multiagent equilibria with limited agents. 2002.

[Dreef *et al.*, 2002] M. Dreef, P. Borm, and B. van der Genugten. On strategy and relative skill in poker. Technical report, 2002.

[Fudenberg and Tirole, 1991] D. Fudenberg and J. Tirole. *Game Theory*. The MIT Press, 1991.

[Gilpin and Sandholm, 2008] A. Gilpin and T. Sandholm. Solving two-person zero-sum repeated games of incomplete information. *Proceedings of AAMAS-08*, 2008.

[Hart, 1985] Sergiu Hart. Nonzero-sum two-person repeated games with incomplete information. *Mathematics of Operations Research*, 10(1):pp. 117–153, 1985.

[Kuhn, 1953] Harold W. Kuhn. Extensive games and the problem of information. *Annals of Mathematics Studies*, 28, 1953.

[McKelvey and Palfrey, 1995] Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38, July 1995.

[Myerson, 1978] R. Myerson. Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, 7:73–80, 1978.

[Selten, 1975] R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4(1):25–55, 1975.

[Shalev, 1994] Jonathan Shalev. Nonzero-sum two-person repeated games with incomplete information and known-own payoffs. *Games and Economic Behavior*, 7(2):246–259, September 1994.

[Shoham and Leyton-Brown, 2009] Y. Shoham and K. Leyton-Brown. *Multiagent Systems: algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2009.

[von Neumann, 1928] J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100:295–320, 1928.