On Reputation in Game Theory Application on Online Settings

Karl Aberer, Zoran Despotovic
Distributed Information Systems Laboratory (LSIR)
Institute of Core Computing Sciences (IIF)
Swiss Federal Institute of Technology (EPFL)
Lausanne, Switzerland

1 Introduction

The phenomenon of reputation has been extensively studied in economics, game theory in particular. The game theoretic framework for analysing reputation is that of repeated games in which some players are uncertain about the payoff structures of their opponents. There are two keyterms in this informal definition: repeated game and uncertainty. Both of them are well in accordance with informal understanding of the notion of reputation. The necessity of repetition is rather obvious - it makes sense to build reputation only if it can influence future returns. If all interactions are one time deals then reputation is not needed because no one will happen to know in the future what you did in the past. Uncertainty is just a little bit less obvious: one who is *locked* into specific type of behavior, without any other options to choose from, cannot have reputation for that specific type. Rather, he or she *is* of that type with certainty.

In this text we present how game theory operationalizes this informal view on reputation and discuss how the developed models can be applied in open electronic communities to promote trustworthy behavior. Sections 2 and 3 lay out the ground for understanding complex reputation models by descibing basic game-theoretic concepts including games with incomplete information and repeated games. Section 4 introduces game-theoretic reputation models on the example of the well known chain-store game, while Section 5 presents generalizations of the observations made on this example. These two sections deal with so called models with perfect monitoring in which the players are fully informed about the past plays of their opponents. Sections 6 and 7 discuss more recently developed models in which the players observe only imperfect signals of the past play. In Section 6 the signals are public, common for all players, while Section 7 deals with the case of privately observed, different, signals. We conclude in Section 8 with a discussion on possible ways to apply this game-theoretic modeling on online settings.

2 Game Theory Basics

Game theory describes mathematical models of conflicting and cooperative interactions between rational, utility maximizing, decision makers (players in the parlance of game theory). The presence of interaction is the most important ingredient of this definition - the utilities of the players are affected not only by their own strategic choices but also by those of all other players as well.

In this text we will concentrate on so called *strategic games*, in which players choose their actions simultaneously, at the same time. This concept can be formally defined as follows.

Definition 2.1. A *strategic game* consists of:

- a set of players $N = \{1, 2, \ldots, n\}$,
- for each player i, a pure strategy (or action) set A_i and
- for each player i, a utility function $u_i : A \to \mathbb{R}$, where $A = \times_{i \in N} A_i$.

□

The definition is illustrated in the following example.

Example 2.1. Consider a conflicting situation two hunters might face. Assume that they have the possibilities to hunt a stag (together) or hares (separately). Thus we have that the set of players is two-element set $N = \{1, 2\}$ and that the players' strategy sets are $A_1 = A_2 = \{\mathsf{Stag}, \mathsf{Hare}\}$. To have a completely defined game we need also to define the players' utility functions on the set $A_1 \times A_2$. They can be neatly presented by a table, as done in Table 1. Each cell in the table contains two entries corresponding to the utilites players 1 and 2 respectively receive when the combination of actions corresponding to the cell has been chosen.

		Player 2	
		Stag	Hare
Player 1	Stag	2, 2	0, 1
	Hare	1, 0	1, 1

Table 1: A two-player Stag Hunt game.

□

The above definition also introduces some notational conventions we will be using. Apart from the symbols given there we will use the symbol $\Delta(B)$ to denote the set of all probability distribution over the set B. For any player

i the set $\mathscr{A}_i = \Delta(A_i)$ will be called the set of player i's *mixed* strategies. Thus, mixed strategies of player i are probability distributions over his pure strategies A_i. (Where it does not make any confusion we will use the notion strategy to denote both mixed and pure strategies.) The set of all mixed strategy combinations, $\times_{i \in N} \Delta(A_i)$ will be denoted by $\mathscr{A}$. Any mixed strategy profile $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathscr{A}$ induces a probability distribution over the set of all pure strategy profiles A in the obvious way. For any such mixed strategy profile α we define $u_i(\alpha)$ as the expected value of player i's payoff with respect to this distribution. Also, any strategy profile $(a_1, \ldots, a_i, \ldots, a_n)$ will be denoted by (a_i, a_{-i}) when we want to emhasize that player i uses action a_i. Generally, any combination of objects of interest in which player i do not participate will have $-i$ in the subscript. The set of all players without player i will be denoted by $N - i$.

We now ask the question: how will a game such as the one just shown be actually played. Or, what strategies maximize the players' utilities? These questions are central to game theory and the solution concept game theorists prescribe is that of *Nash equlibrium*. Informally, the concept is made of two components: 1) players must form beliefs about other players' strategies and then, based on these, choose their own strategies optimally and 2) the players' beliefs must be consistent in the sense that every player's belief about the other players' actions is correct.

The idea is formalized in the following definition.

Definition 2.2. A strategy profile $\alpha^* = (\alpha_1^*, \ldots, \alpha_n^*) \in \mathscr{A}$ is a Nash equlibrium if no player can gain by unilateraly deviating. So, denoting $\alpha^* = (\alpha_i^*, \alpha_{-i}^*)$, we have that for any player i and any of her strategies $\alpha_i \neq \alpha_i^*$ the following must be satisfied:

$$u_i(\alpha_i^*, \alpha_{-i}^*) \geq u_i(\alpha_i, \alpha_{-i}^*) \tag{2.1}$$

□

By inspecting Table 1 we can see that the profiles (Stag, Stag) and (Hare, Hare) are Nash equlibria of the stag hunt game. For instance, given that one player plays Stag the other player cannot increase his utility by switching to any randomization between Stag and Hare that puts a positive probability on Hare. But these two equilibria (in pure strategies, obviously) are not the only ones. There is another involving mixed strategies, in which both players randomize between their two pure strategies with probabilities of 0.5.

The question of the equilibria existence is resolved by the following theorem.

Theorem 2.1. For any finite game there is at least one Nash equilibrium in $\mathscr{A} = \times_{i \in N} \mathscr{A}_i$.

□

The theorem is an application of the fixed point theory for point-to-set correspondences. Border (1985) presents a good introduction to this subject.

2.1 Bayesian Games

In many games of interest different players may have different information about some important aspects of the game, including a special case in which some players know more than others and may use this informational asymmetry to get better payoffs. To model such situations the concept of *Bayesian games* (or games with incomplete information), introduced by Harsanyi (1967–68), has been widely adopted. Informally, central to Bayesian games are *types* of the players by which the players' private information is modeled. Each player learns his own type at the beginning but the types of other players remain unknown to him throughout the game. The utility functions are also different than before - they specify ordinal payoffs for combinations of chosen strategies and realized types rather than strategies only. Another important assumption is that the types are drawn from a common prior probability distribution, known to and agreed upon by all the players, so that every player can derive the probability distributions of the combinations of the other players' types given any of his own types. Then, the goal of every player is to maximize his payoff for any of his types. As before, in an equlibrium any player must form a belief about the other players' strategies given his knowledge of the types' distributions and then play optimally. Of course, all beliefs must be consistent. The formal definitions follow.

Definition 2.3. A *Bayesian game* consists of a set of players $N = \{1, 2, \ldots, n\}$ and for each player i:

- a set of possible actions A_i,
- a set of possible types T_i,
- a probability function $p_i : T_i \to \Delta(T_{-i})$, where $T_{-i} = \times_{j \in N-i} T_j$ and
- a utility function $u_i : A \times T \to \mathbb{R}$, where $A = \times_{i \in N} A_i$ and $T = \times_{i \in N} T_i$.

□

Definition 2.4. Any strategy profile $\sigma^* \in \times_{i \in N} \times_{t_i \in T_i} \Delta(A_i)$ is a Bayesian equlibrium if for any type t_i of any player i mixed strategy $\sigma^*(a_i|t_i)$ optimizes player i's expected payoff where the expectation is taken over all combinations of types of the other players.

□

Example 2.2. Auctions present a typical example of Bayesian games. In this example we will describe the first price sealed-bid auction. Let us assume that n bidders are competing to buy an auctioned item by submitting sealed bids. The owner of the highest bid wins and pays the price of his bid. Assume also that any bidder has a private valuation of the item, denoted v_i, and that each player's valuation is independent of those of the other players. However, the distribution from which the valuations are drawn are known. Here we assume the unifom distribution on the interval $[0,1]$. So, it should be obvious that this makes a Bayesian game in which the item valuations are the players' types and all distributions p_i for any player i are uniform on $[0,1]$. The utility functions are such that for any player i and any valuation-bid tuple $(v,b) = (v_1,\ldots,v_n,b_1,\ldots,b_n)$

$$u_i(v,b) = \begin{cases} v_i - b_i & \text{if } b_i = \max(b_1,\ldots,b_n) \\ 0 & \text{otherwise.} \end{cases}$$

It can be shown (a proof can be found in almost any introductory text on auctions, McAfee and McMillan (1987) for instance) that in this setting the optimal bid for any bidder i with valuation v_i is $\frac{n-1}{n}v_i$. In other words, the equlibrium is made of the set of n functions $B(v) = \frac{n-1}{n}v$ for n bidders and $v \in [0,1]$.

□

3 Repeated Games

To model repeated interactions game theorists use the concept of repeated games in which the same stage game is repeated finite or infinite number of times whereby the sets of players who participate in the stages can vary from stage to stage. In this section our primary concern will be infinitely repeated games with discounting, in which the stage game is repeated infinitely many times and the players discount their future payoffs as compared to those received at present. The main goal of the section will be to show that repeated play can differ greatly from one-shot encounters in the sense that it can allow for a whole range of equilibria which are not normally found in the constituent, one-shot games. Building on the results we present here, we will show in the later sections how reputation effects (enabled by adding uncertainties in the Bayesian sense we just presented) may help certain players to narrow down these equilibrium sets and pick those equilibria they prefer the most.

To define a repeated game and its equilibria we need to define the players' strategy sets and payoffs for the entire repeated game given the strategies and payoffs of its constituent stage game. Therefore, assume that the stage game is an n-player strategic form game and that the action set of each player i, A_i, is finite. The stage game payoffs are defined in the usual way, as maps $u_i : A \to \mathbb{R}$, where $A = \times_{i\in N} A_i$. Let also $\mathscr{A}_i$ be the space of all mixed strategies of player i. Assuming that the players observe each other's realized pure strategies at the end of each stage and are aware of them when choosing the next stage strategies we can proceed as follows: let $a^t = (a_1^t,\ldots,a_n^t)$ be the pure strategy profile relized in period t and $h^t = (a^0,\ldots,a^{t-1})$ the history of these profiles for all periods before t. Let, finally, H^t denote the set of all such period-t histories. Then a period-t mixed strategy of player i in the repeated game is any mapping $\sigma_i^t : H^t \to \mathscr{A}_i$. A mixed strategy player i for the whole repeated game is then a sequence of such maps $\sigma_i = \{\sigma_i^t\}_{t=1}^{\infty}$.

Now, there is one more clarification we have to make in order to define the repeated game payoffs and equilibria. Namely, in all stages of the game players receive some payoffs and the problem is how to compare, possibly infinite, sequences of these stage game payoffs. There are three criteria adopted in the literature: time-average, overtaking and discounting criterion. For the discounting criterion is the most widely accepted and most closely approximates the real-world situations we will in this text concentrate exclusively on it. So if u_i^t is the payoff player i receives in the period t then the δ-discounted $(0 \leq \delta < 1)$ payoff of this player is defined as:

$$g_i = (1 - \delta) \sum_{t=0}^{\infty} \delta^t u_i^t \tag{3.1}$$

The reasoning behind this definition is that the players see future gains as less valuable than those of the present or that there is uncertainty regarding the ending time of the game with $1 - \delta$ being the probability that the next stage will be the last. For obvious reasons, a repeated game in which the players discount their future payoffs with a common discounting factor δ will be denoted $G(\delta)$.

Any strategy profile $\sigma = (\{\sigma_1^t\}_{t=1}^{\infty}, \dots, \{\sigma_N^t\}_{t=1}^{\infty})$ induces a probability distribution over the set of all infinite histories. We define the players' payoffs in the repeated game as the expected values of δ-discounted payoffs the players receive when the paths of play follow each of these histories. With this clarification we finally define that a Nash equlibrium of the repeated game $G(\delta)$ is every strategy profile $(\{\sigma_1^t\}_{t=1}^{\infty}, \dots, \{\sigma_N^t\}_{t=1}^{\infty})$ such that no player can gain by switching to a different strategy given that all other players implement it.

The central problem in the theory of repeated games, investigated by many authors, is what payoff sets can be supported in equilibria of the repeated game given the structure of its constituent, stage game. Intuitively, because the players can condition their future play on the past play of their opponents and retaliate if the opponents do not play in a specific way, many payoff allocations, not supportable in the one shot game, might become supported in an equilibrium of the repeated game provided the future is not discounted too much. The theorems formalizing this intuition are known in the literature as *folk theorems*. As seen from the perspective of engineering of online reputation mechanisms, they are important as the starting points determining what outcomes are feasible for a given mechanism. If socially desirable outcomes are not among the feasible ones then the mechanism must be redesigned as to include them.

We will state here two folk theorems corresponding to two types of repeated interactions: one with all players participating in all stages of the game (hereafter such players will be termed long-run) and one in which one player is long-run while his opponents play only once (short-run hereafter).

We first cite the result for the setting with all players being long-run. To state it precisely we need to introduce several notions. We define the *minmax value* of player i to be

$$\underline{v}_i = \min_{\alpha_{-i} \in \times_{j \in N-i} \Delta(A_j)} \max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \alpha_{-i}). \tag{3.2}$$

Thus, player i's minmax value is the lowest payoff he can achive in the stage game provided he correctly foresees the choice of actions of his opponents. Any payoff $v_i > \underline{v}_i$ is called individually rational for player i.

Further, putting $u(a) = (u_1(a), \ldots, u_n(a))$, we can define the set of feasible payoffs as

$$V = \text{convex hull}\{v | \exists a \in A \text{ s.t. } u(a) = v\}. \tag{3.3}$$

Theorem 3.1 (Fudenberg and Maskin (1986)). For any feasible payoff allocation v that is individually rational for all players there is a $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a Nash equilibrium of the repeated game $G(\delta)$ with payoffs v.

□

Example 3.1. As an example, let us apply this theorem to the Prisoner's Dilemma game, shown in Table 2. The only equilibrium of the game is for the both players to cheat, which implies that the socially desirable outcome, in which the players cooperate, does not constitute an equilibrium. But the things change considerably in the game is repeated. Namely, it is easy to see that the minmax values of both players are $\underline{v}_1 = \underline{v}_2 = 1$ and that the payoff allocation $(5, 5)$ is in the set V. Thus, we can conclude that the cooperative outcome with both players cooperating in every stage of the repeated game is possible in the long term interaction provided the players are sufficiently patient (their discount factors are suficiently close to 1). It is also easy to construct such a strategy profile that results in this cooperative equlibrium. For instance, if the players implement the strategy "cooperate as long as no player cheated in the past stages, otherwise cheat" then no player has incentive to deviate if his short-term gain from the deviation is less then his long-term gain from cooperation. This will be the case if receiving 5 utils in every stage is better than receiving 6 today and 1 in all future stages, or if $5 \geq (1-\delta)6 + \delta$, which implies $\delta \geq \frac{1}{5}$. For an elaborate exposition of this phenomenon and its empirical confirmations we refer to Axelrod (1984).

	Cooperate	Cheat
Cooperate	5, 5	0, 6
Cheat	6, 0	1, 1

Table 2: Payoff Matrix of the Prisoner's Dilemma Game

□

Of particular importance for application to online world is the repeated game model with one long-run player facing a (possibly infinite) sequence of short-run players each of whom plays the stage game only once as opposed to the long-run player who stays in the game till its end. This model closely approximates a typical online setting, P2P in particular, in which a large number of participants makes multiple interactions between the same players highly improbable.

The question raised above, what payoff allocations can be achieved in an equilibrium of the repeated game, is again in the focus of our attention. The problem is a little different now because the short-run players are unconcerned about the future and will always play their stage game best responses to whatever strategies are played by their opponents. To formulate the results clearly we will again need a few definitions (to keep the discussion simple we state the definitions and the results for the case of a two players stage game in which player 1 is long-run while player 2's are short-run).

Let

$$B : \mathscr{A}_1 \rightrightarrows \mathscr{A}_2 \tag{3.4}$$

be the correspodence that maps the long-run player's mixed strategies to the best response strategies of the short-run player. So for any pair $(\alpha_1, B(\alpha_1))$ from $\mathscr{A}_1 \times \mathscr{A}_2$ we have that $B(\alpha_1)$ is player 2's best response strategy to α_1. Let the set of all such pairs be denoted by graph(B). The long-run player's minmax value must be also redefined as to include the constraint that the short-run players play the stage game best responses. So the minimax value of the long-run player is given by:

$$\underline{v}_1 = \min_{\alpha \in \text{ graph}(B)} \max_{\sigma_1} u_1(\sigma_1, \alpha_{-1}). \tag{3.5}$$

Before we are ready to state the folk theorem for this setting, we need just one simple definition: for any strategy $\sigma_1 \in \Delta(A_1)$, supp(σ_1) is defined as the set of all $a_1 \in A_1$ on which σ_1 puts a non-zero probability.

Theorem 3.2 (Fudenberg, Kreps, and Maskin (1990)). Let

$$\bar{v}_1 = \max_{(\sigma_1, \sigma_{-1}) \in \text{ graph}(B)} \min_{s_1 \in \text{supp}(\sigma_1)} u_1(s_1, \sigma_{-1}). \tag{3.6}$$

Then for any $v_1 \in (\underline{v}_1, \bar{v}_1)$ there exists a $\delta' < 1$ such that for all $\delta \in (\delta', 1)$, there is a subgame perfect equilibrium in which player 1's discounted payoff is v_1. For no δ is there an equilibrium where player 1's payoff exceeds $\bar{v}_1$.

□

For completness, we should add that an underlying assumption of this theorem is that only pure actions realized in the past stages are observable to the players. If, on the contrary, the past mixtures of the players are known then a folk theorem similar to Theorem 3.1 obtains.

Example 3.2. It is easy to apply this theorem to the repeated Prisoner's Dilemma game with one long-run and one short-run player. Namely, it is easy to see that graph(B) can include only strategy profiles in which player 2 plays Cheat because this strategy is player 2's best response strategy no matter what player 1 does. Thus, we have $\underline{v}_1 = \bar{v}_1 = 1$, which brings us to the conclusion that the only possible equilibrium of this version of the repeated Prisoner's Dilemma is for the both players to cheat in all stages and, therefore, no cooperative equilibrium can be attained.

□

	In	Out
Acquiesce	1, 2	3, 0
Fight	0, -1	3, 0

Table 3: Payoff Matrix of the Chain Store Game

Example 3.3. Table 3 shows a simultaneous move version of the celebrated chain store game, studied by Selten (1978).[1] Assume that the the row player is a firm already established in a market (monopolist hereafter) and that the column player is another firm that decides on entering the market or staying out (entrant hereafter). The monopolist can choose to fight the entry (say, it may opt to a sharp price cut) or to acquiesce (to share the market peacefully). The payoffs of the two firms are as shown in Table 3.

To determine graph(B) we can proceed as follows. If the monopolist randomizes between its two strategies with probability p put on $\mathsf{Acquiesce}$ then playing In would bring the entrant the expected payoff of $3p - 1$ while Out always brings 0. So, the entrant's best response is Out if $p \leq \frac{1}{3}$ and In if $p > \frac{1}{3}$. Put another way we have:

$$\text{graph}(B) = \begin{cases} (p \cdot \mathsf{Acquiesce} + (1-p) \cdot \mathsf{Fight}, \mathsf{Out}) & \text{if } 0 \leq p \leq \frac{1}{3} \\ (p \cdot \mathsf{Acquiesce} + (1-p) \cdot \mathsf{Fight}, \mathsf{In}) & \text{if } \frac{1}{3} < p \leq 1 \end{cases}$$

It is easy to see now that the maximum in (3.6) is achieved for any profile $\sigma = (p \cdot \mathsf{Acquiesce} + (1-p) \cdot \mathsf{Fight}, \mathsf{Out})$ with $p \leq \frac{1}{3}$ (both minima equal 3, which is the monopolist's maximum payoff in the stage game). Therefore, we can conclude that in the repeated game in which the monopolist is long-run while the entrants are short-run the equlibrium in which the monopolist receives the payoff 3, and consequently, in which an entry never ocurrs, is possible. However, it can be also seen that $\underline{v}_1 = 1$, meaning that many other equilibria, in which player 1's payoffs range from 1 to 3, are also possible. As we will see shortly, reputation is exactly what can help player 1 narrow down this set and get payoffs arbitrarily close to the best possible one (3 in this example).

□

[1] The sequantial move version of the game is shown in Figure 1 on page 11. To keep our discussion simple we will not elaborate on the details of this class of games, also known as extensive form games.

4 Uncertainty and Reputation Formation - Perfect Monitoring Case

To recapitulate, we started with an assertion that for reputation formation repeated interactions and interactants' uncertainties are necessary. Then we described two game theoretic concepts, Bayesian and repeated games, that are used to model each of these two. In particular, we anylised the repeated game settings with one long-run player facing a sequence of short-run opponents because this model is the most appropriate for online settings. The most important conclusion we could draw from this analysis was that a multitude of equilibria is possible in repeated games owing to the possibility for the players to punish in the future stages any past "misbehaviors" of their opponents.

In this section we add uncertainty to the repeated game setting and describe how game theory joins these two concepts to model reputation effects. Our main objective will be to show how the long-run player can develop reputation for being of a certain type and thus make credible the intuition that he may always pick the equlibrium he prefers the most.

Monopolist weak

	In	Out
Acquiesce	0, b	a, 0
Fight	-1, b -1	a, 0

Monopolist strong

	In	Out
Acquiesce	-1, b	a, 0
Fight	0, b-1	a, 0

a>1
0< b <1

Table 4: Payoff Matrix of the Bayesian Chain Store Game

The chain-store game from Example 3.3 and its Bayesian variation from Table 4 will serve as the running examples for this section. The setting we will consider will be that of one long-run monopolist playing against many short-run entrants. An important assumption of this section is that at any stage of the repeated game the players are perfectly informed about all previous play. To be precise, the term "perfectly" means that at any stage they know all actions taken by all players at all past stages.

The chain-store game from Table 3 has three equlibria: one in which the monopolist acquiesces and the entrant enters, one in which the monopolist fights and the entrant stays out, and one in which the monopolist fights with probability 2/3 and the entrant stays out. If we change the definition of the game in such a way that the order of moves matters and the entrant moves first we get a game tree as in Figure 1. This is an example of so called extensive form (sequential move) games. Generally speaking, this class of games needs a different sort of analysis than the class of strategic games, but in this simple case the equilibrium sets of the two games are the same. The reason why we present this sequential move version is to show that if the order of moves matters (and we will adopt this in the analysis below) then only the first one among the equilibria given above is plausible: if presented with an entry the monopolist must acquiesce, the threat of fighting an entry when an entry ocurrs is not really credible. (In the rest of this text we will not pay much attention to this class of games and therefore we opt not to show the details related to them.)

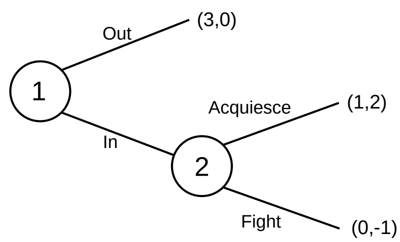

Figure 1: Sequential Move Version of the Chain-Store Game

Let us see what will happen if this game is repeated finitely many times. An intuition says that it might make sense for the monopolist to fight some early entries in an attempt to convince the future entrants that he is "tough" (the precise meaning of this term will become clear shortly) and will fight future entries as well. Then from a certain point on no entry would be expected to ocurr and the losses incurred by fighting early entries would be offset by not sharing the market in the later stages. Of course, this intuition might be valid only if the monopolist does not discount the future payoffs too much (its δ is sufficiently close to 1) or the number of stages is sufficiently high. But this intuition is wrong. Namely, the last entrant knows that there is no more entries for the monopolist to deter and will, therefore, enter. It is important to see that this will happen in the last stage no matter what was the prevous play. The second last entrant will predict this and conclude that the monopolist can by no means deter entry in the last stage and, therefore, will not fight the second last one. Thus, the second last entrant will enter too. So if we unfold the whole sequence of play from its end backwards to the begining we reach the conclusion that all the entrants will enter. This reasoning, seemingly counterintuitive, is referred to in the literature as "chain-store paradox" (Selten (1978)).

The missing ingredient in the model, that would justify the mentioned intuition, is the entrants' uncertainty with respect to the payoff structure of the monopolist. Put another way, the monopolist cannot develop reputation for "toughness" when it cannot be tough whatsover. As soon as the entrants assign even a small prior probability to the event that the monopolist benefits from fighting entry (it is though), an equilbrium in which entries do not ocurr but in several last stages of the entire game can be constructed. This is exactly the approach taken by Kreps and Wilson (1982) and Milgrom and Roberts (1982), the two works that introduced reputation in game theory.

Table 4 shows the Bayesian version of the chain-store game analysed by Kreps and Wilson (1982). It can be seen from the table that the monopolist has two types: it can be weak (the payoffs are in this case just slightly different than in the original game given above) and strong, in which case it prefers fighting the entry than acommodating it. The entrants initially assign a non-zero probability to the monopolist being strong (let us denote it by p_0), while there is no uncertainty regarding their payoffs. This stage game is repeated and after any stage k the probability that the monopolist is strong is updated for the next stage $k+1$ given the monopolist's play at stage k according to Bayes' rule: if the the monopolist fought the entry in stage k then $p_{k+1} \geq p_k$, otherwise $p_{k+1} = 0$ because the monopolist then reveals that he is weak. An equilibrium, as usual, implies that: 1) at any stage and for any history of the previous moves the monopolist's strategy is a best response to the entrants' strategy and 2) for each entrant i and each history of the play prior to stage i entrant i's strategy is a best response to the monopolist's strategy given that the monopolist is strong with the corresponding probability.

Instead of stating the equilibrium strategies formally we will outline a typical equlibrium play. Namely, at the start of play the entrants know that the monopolist, having enough stages left for benefiting from building reputation, will fight entries and thus they choose to stay out. The play proceeds this way until the last several stages, when the entrants get enough incentive to challenge the monopolist and when entries ocurr either with certainty or with some probability less than one.

A remarkable property of this equilibrium is that the number of these "last stages" in which entries might happen is independent of the total number of the stages played (it depends only on the inital probability of the monopolist being weak, p_0). This implies that the total number of stages grows the average payoff per stage received by the monopolist becomes arbitrarily close to a, which is the best stage game payoff the monopolist would want to get.

There is another very interesting property of this equilibrium. If the entrants's beliefs are updated in such a way that $p_{k+1} \geq p_k$ if entry k was fought (Kreps and Wilson originally used the term "plausible") then this equilibrium is unique (there are many other equlibria that do not satisfy this requirement). This is particularly important if we know that the problem of the equlibrium selection (or so called focal points) when a game (repeated or one-shot) has multiple equilibria is still unresolved question in game theory (Myerson (1991) presents a good discussion on this subject). The uniqueness and plausibility make us believe that this equlibium will be picked in an actual play of this game so that even this problem of equilibrium selection is eliminated.

5 Perfect Monitoring – General Results

The result of Kreps and Wilson, we just described, generated quite some interest in studying reputation formation and the very next problem game theorists attacked was its generalization through identification and characterization of the necessary ingredients of the model that brought to such a result.

Fudenberg and Levine (1989) considered a general setting with a patient (long-run) player playing the stage game G against an infinite seuence of short-run players. They define "Stackelberg" outcome in G as the long-run player's most prefered pure strategy profile given that the short-run players always choose their best responses to the strategy of the long-run player. Formally, let

$$g_1^* = \max_{s_1 \in A_1} \min_{\sigma_2 \in B(s_1)} u_1(s_1, \sigma_2). \tag{5.1}$$

and let s_1^* be the strategy achieving the maximum in (5.1). Then g_1^* is called the long-run player's Stackelberg payoff and s_1^* his Stackelberg strategy.

In the game from Table 3.3, for example, Fight is the long-run player's Stackelberg strategy and the payoff 3, corresponding to the strategy profile (Fight, Out), is his Stackelberg payoff.

The model further assumes that the player 1's payoffs are associated with his types (as in any Bayesian game), which are drawn from a set $\Omega = \{\omega_0, \omega_1, \ldots\}$ with a prior probability distribution μ defined on it. Crucial to the model is the assumption that there must be at least two types of the long-run player: one corresponding to the payoffs of the original stage game G (often referred to as "sane" type, denote it by ω_0) and one having the Stackelberg strategy as its dominant strategy in the repeated game ("crazy" type, ω_1). Let us also assume that their prior probabilities are $\mu(\omega_0) = \mu_0$ and $\mu(\omega_1) = \mu_1$. The following theorem asserts that in this setting the least payoff of the long-run player in any Nash equilibrium of the repeated Bayesian game $G(\delta, \mu)$ equals almost g_1^*.

Theorem 5.1 (Fudenberg and Levine (1989)). Let $\mu(\omega_0) > 0$ and $\mu(\omega_1) > 0$ and let $\min g_1$ denote the minimum payoff of the long-run player in the stage game. Then there is a constant $k(\mu_1)$ otherwise independent of (Ω, μ), such that the least payoff to the long-run player of type ω_0 is

$$\underline{V}(\delta, \mu, \omega_0) \geq \delta^{k(\mu_1)} g_1^* + (1 - \delta^{k(\mu_1)}) \min g_1 \tag{5.2}$$

□

At the heart of Theorem 5.1 is the assertion that, given any $0 \leq \pi < 1$ and given that the long-run player always plays his Stackelberg strategy s_1^*, there is an upper bound on the number of occasions in which the short-run players assign a probability of the long-run player being of the type ω_1 smaller than π. Fudenberg and Levine (1989) show that $k = \log \mu_1 / \log \pi$ (where μ_1 is the probability of the long-run player being of Stackelberg type) can be used as this bound.

It is easy to see from (5.2) that when δ grows the bound grows too and, in the limit case, $\lim_{\delta \to 1} \underline{V}(\delta, \mu, \omega_0) = g_1^*$. Thus, the bound is higher for more patient players and the lowest Nash equlibrium payoff of the fully patient players is arbitrarily close to their Stacklberg payoff.

Example 5.1. Let us apply these results to the chain-store game from Table 3.3. As we already saw, if the probability of the monopolist fighting the entry exceeds 2/3 then the short-run player will strictly prefer to stay out. In other words, we can set $\pi = 2/3$ and compute that, with $\mu_1 = 0.05$ for instance,

$k = \log \mu_1 / \log \pi = (\log 0.04)/(\log 2/3) \approx 8$. Thus, the short-run players will enter at most 8 times. In the worst case, if these entries happen at the very begining of the game and the long-run player receives the lowest payoff in each of these stages (which equals 0) then we see that the above bound becomes $\underline{V}(\delta, \mu, \omega_0) = \delta^8 g_1^*$.

□

6 Imperfect Public Monitoring

As we have seen, an important assumption of the previous section was that the players were fully informed about all past action choices prior to choosing their own actions. In this section we will drop this assumption and show what happens in an even more realistic scenario - when past actions of the players are not observed but, instead, only their imperfect signals. As before, we remain focused on repeated games with one long-run player and many short-run opponents.

The model can be defined as follows. The long-run player (player 1) plays a stage game against an infinite sequence of one-shot players 2's. As usual, the strategy sets in the stage game will be denoted by A_1 and A_2, while $\mathscr{A}_1$ and $\mathscr{A}_2$ will denote the corresponding mixed strategy spaces. As compared to the models of perfect monitoring, the stage game actions played in the past rounds are not observable at any period of the game. Instead, what is observable now are only stochastic outcomes y corresponding to the actually implemented action profiles. These outcomes are drown from a finite set Y according to a probability distribution $\rho(\cdot|a)$ for any $a \in A_1 \times A_2$. They are commonly observed by all the players (the monitoring is, thus, public) and the same outcome can ocurr after many different action profiles implemented (the monitoring is, therefore, imperfect). Finally, we are assuming that the probability distributions ρ themselves are common knowledge among the players.

The model further assumes that the players' payoffs in any round t depend only on the observed signal and their corresponding strategies implemented in round t. (Put another way, the payoffs depend on the actually implemented profile $a = (a_1, a_2)$ only through its influence on the distribution ρ.) Thus, the long-run player's payoffs is any function $u_1 : Y \times A_1 \to \mathbb{R}$, while each short-run player's payoff is a function $u_2 : Y \times A_2 \to \mathbb{R}$. The ex ante stage game payoffs when the action profile $a = (a_1, a_2)$ has been actually implemented are $v_1(a_1, a_2) = \sum_{y \in Y} u_1(y, a_1)\rho(y|a)$ and $v_2(a_1, a_2) = \sum_{y \in Y} u_2(y, a_2)\rho(y|a)$.

The players' strategies can be defined as follows. Assume that at any time t the history of previously observed signals is $(y^0, \ldots, y^{t-1}) \in H^{t-1}$, the history of the long-run player's actions is $(a_1^0, \ldots, a_1^{t-1})$ and that the history of the short-run players' actions is $(a_2^0, \ldots, a_2^{t-1})$. So the knowledge the long-run player has on the play prior to time t is $h_1^{t-1} = ((a_1^0, y^0), \ldots, (a_1^{t-1}, y^{t-1})) \in H_1^{t-1} \equiv (A_1 \times Y)^t$ while that of the period-t short-run player (assuming that this short-run player knows the action choices of the prevous player 2's) becomes $h_2^{t-1} = ((a_2^0, y^0), \ldots, (a_2^{t-1}, y^{t-1})) \in H_2^{t-1} \equiv (A_2 \times Y)^t$. With this notation we define player 1 strategy as any map $\sigma_1 : \cup_{t=0}^{\infty} H_1^t :\to \mathscr{A}_1$ while player 2's strategy can be defined as a map $\sigma_2 : \cup_{t=0}^{\infty} H_2^t :\rightarrow \mathscr{A}_2$. (For notational convenience, we treat the short-run players as one player. Of course this does not impose any problem if we define the equilibria in such a way that σ_2^t is a best response to σ_1 for each t.)

Consider now a strategy profile $\sigma^* = (\sigma_1^*, \sigma_2^*)$. Player 2's strategy σ_2^* is a best response to σ_1^* if for any t, any $h_2^{t-1} \in H_2^{t-1}$ and any $j \in A_2$:

$$E_{\sigma^*}[v_2(a_1^t, a_2^t)|h_2^{t-1}] \geq E_{\sigma^*}[v_2(a_1^t, j)|h_2^{t-1}].$$

Now, for a strategy profile $\sigma^* = (\sigma_1^*, \sigma_2^*)$ to be a Nash equlibrium it is necessary that, apart from the requirement that σ_2^* be a best response to σ_1^*, σ_1^* must maximize the long-run player's expected utility. This will be achieved if for all σ_1 the folloving inequality holds:

$$E_{\sigma^*}[(1-\delta)\sum_{s=0}^{\infty} \delta^s v_1(a_1^s, a_2^s)] \geq E_{(\sigma_1, \sigma_2^*)}[(1-\delta)\sum_{s=0}^{\infty} \delta^s v_1(a_1^s, a_2^s)]. \tag{6.1}$$

To study reputation effects we allow for the possibility for the long-run player to have one or more additional, commitment types each of which prefers to play a specific (possibly mixed) stage game strategy. The definition of a Nash equilibrium is just slightly different in the case of this perturbed, incomplete information game: assuming (for simpler exposition) that there is only one commitment type who always plays strategy $\hat{\sigma}_1$ we can say that in any equilibrium $\sigma^* = (\sigma_1^*, \sigma_2^*)$, σ_2^* must maximize

$$p_t E_{(\hat{\sigma}_1, \sigma_2)}[v_2(a_1^t, a_2^t)|h_2^{t-1}] + (1-p_t) E_{(\sigma_1^*, \sigma_2)}[v_2(a_1^t, a_2^t)|h_2^{t-1}]$$

for any t and any history h_2^{t-1}. In this expression p_t is period-t player 2's posterior belief that the long-run player is the commitment type. It is assumed that the prior p_0 is known to all the players. The long-run player's equilibrium strategy σ_1^* maximizes the same objective function (6.1).

Cripps, Mailath, and Samuelson (2002) consider exactly the setting we just described and show that (under some mild conditions) reputations for non-credible behavior *cannot* be permanently maintained. We will state this result precisely.

The concept of strategy that is *never an equilibrium strategy in the long-run* plays the central role. So player 1's strategy $\bar{\sigma}_1$ is never an equilibrium strategy in the long-run if there exists T such that for all histories h_2^t, all player 2's best responses to $\bar{\sigma}_1$ and all $t \geq T$ player 1 always has some profitable deviation.

The authors further introduce several assumptions regarding the distribution ρ of the signals:

Assumption 1 (Full Support). $\rho(y|a) > 0$ for all $a \in A_1 \times A_2$.

□

Assumption 2 (Identification).

- For all $a_1 \in A_1$ there are $|A_2|$ linearly independent columns of the matrix $[\rho(y|(a_1, a_2))]_{y \in Y, a_2 \in A_2}$.
- For all $a_2 \in A_2$ there are $|A_1|$ linearly independent columns of the matrix $[\rho(y|(a_1, a_2))]_{y \in Y, a_1 \in A_1}$.

□

Assumption 1 simply says that any signal has a positive probability under any pure strategy profile, while assumption 2 requires that the players can statistically identify each other's stage game actions based on a sufficent observations of the signals. (In the spirit of Fudenberg, Levine, and Maskin (1994) this means that all pure action profiles satisfy the individual full rank condition.)

The main result can be now stated as follows.

Theorem 6.1 (Cripps, Mailath, and Samuelson (2002)). Assume that ρ satisfies Assumptions 1 and 2 and that $\hat{\sigma}_1$ is a commitment strategy of the long-run player which is pure and never an equilibrium strategy in the long-run. Then in any Nash equilibrium $\tilde{\sigma}$, $p_t \to 0$ $P_{\tilde{\sigma}}$-almost surely.

□

So how to interpret this result? Quite generally, there is an important difference between perfect and imperfect monitoring. Namely, as we saw in Section 5 when monitoring is perfect reputational equilibria imply that the long-run player repeatedly plays his Stackelberg strategy while the long-run players' beliefs converge to the Stackelberg type and the long-run player's payoff to his Stackelberg payoff. On the other hand, when monitoring is imperfect then, as the above result says, reputations cannot be maintained forever. Sooner or later short-run players will learn the true type of the long-run player and from that point on the players will play one of the static (stage game) Nash equilibria. According to Dellarocas (2002) an intuitive explanation of this difference is that with perfect monitoring "deviations from the equilibrium strategy reveal the type of the deviator and are punished by a switch to an undesirable equilibrium of the resulting complete-information continuation game. In contrast, when monitoring is imperfect, individual deviations neither completely reveal the deviators type nor trigger punishments. Instead, the long-run convergence of beliefs ensures that eventually any current signal of play has an arbitrarily small effect on the uniformed players beliefs. As a result, a player trying to maintain a reputation ultimately incurs virtually no cost (in terms of altered beliefs) from indulging in a single small deviation from Stackelberg play. But the long-run effect of many such small deviations from the commitment strategy is to drive the equilibrium to full revelation."

Fudenberg and Levine (1992) consider the same scenario and provide lower and upper bounds of the long-run player's payoff in any Nash equilibrium. The result we present here holds for the limit case when the long-run player's discount factor δ is arbitrarily close to 1. We note that the authors provide the bounds for any δ.

For any mixed strategy α_1 of the long-run player the set of generalized best responses $B_0(\alpha_1)$ of the short-run player is defined as containing all $\alpha_2 \in \mathscr{A}_2$ such that α_2 is the best response (in the sense defined before) and there exist some $\alpha_1' \in \mathscr{A}_1$ such that $\rho(\cdot|\alpha_1, \alpha_2) = \rho(\cdot|\alpha_1', \alpha_2)$.

Building on this definition, the authors further define the notions of the least and greatest payoffs when the long-run player is of the "sane" type, $\underline{v}_1$ and $\bar{v}_1$ as:

$$\underline{v}_1 = \sup_{\alpha_1 \in \Delta(A_1)} \inf_{\alpha_2 \in B_0(\alpha_1)} v_1(\alpha_1, \alpha_2) \tag{6.2}$$

$$\bar{v}_1 = \sup_{\alpha_1 \in \Delta(A_1)} \sup_{\alpha_2 \in B_0(\alpha_1)} v_1(\alpha_1, \alpha_2) \tag{6.3}$$

At last, if we denote by $\underline{N}_1(\delta)$ and $\bar{N}_1(\delta)$ the infimum and supremum of the long-run player's payoffs in any Nash equilibrium for his "sane" type and by $\underline{N}_1$ and $\bar{N}_1$ their respective limits when $\delta \to 1$ we obtain:

Theorem 6.2 (Fudenberg and Levine (1992)). If the commitment types have full support then:

$$\underline{v}_1 \leq \underline{N}_1 \leq \bar{N}_1 \leq \bar{v}_1 \tag{6.4}$$

□

7 Imperfect Private Monitoring

A general model of a game with imperfect private monitoring can be defined as follows. Players 1 and 2 (or a sequence of many different player 2's) repeatedly play a stage game with action sets A_1 and A_2. They do not now observe each other's past actions but only their imperfect signals $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$ which are not common. Any strategy profile $a = (a_1, a_2) \in A_1 \times A_2$ generates a probability distribution, $p(\omega|a)$ over the set $\Omega = \Omega_1 \times \Omega_2$ of all combinations of the private signals and these distributions are common knowledge in the game. The stage game payoffs are given by the following sum: $g_i(a) = \sum_{\omega \in \Omega} u_i(a_i, \omega_i) p(\omega|a)$, where $u_i(a_i, \omega_i)$ is player i's realized payoff from the received signal and played action.

Strategies of the players are defined in the usual way, as maps from any history of play a player observes to a probability distribution over his action set. Thus, player i's strategy $(i = 1, 2)$ is any map $\sigma_i : \cup_{t=0}^{\infty} H_i^t :\longrightarrow \mathscr{A}_1$, where $H_i^t \equiv (A_i \times \Omega_i)^t$. With the strategies and payoffs defined, it is easy to define the equilibria of the game. Just as before, an equilibrium is any strategy profile (σ_1^*, σ_2^*) such that no player i can gain by switching to a different strategy σ_i while the other player implements σ_{-i}.

As compared to the model of imperfect public monitoring, there is now only one different detail: the signals the players receive after each stage are not drawn from a common set but from different ones. Put another way, we have $\omega_1 = \omega_2$ when monitoring is public and $\omega_1 \neq \omega_2$ when it is private. Even though the difference between the two models may seem minor at first glance, it is substantial for their tractability: almost perfect understanding of the public monitoring models and an abundance of general results turn into just occasional solutions of particular cases (mostly variations of the Prisoner's Dilemma game) with private monitoring. Dellarocas (2003) presents a notable exception. We will briefly describe this work in Section 8.

Kandori (2002) explains what difficulties arise when monitoring is private and why this model becomes so complicated. He finds the following two reasons the most important: the lack of recursive structure and the need for the players to conduct statistical inference at each stage.

The most part of the literature on the model with imperfect public monitoring considers the case when the players do not condition their play on their private signals (these are played actions) but only on histories of public outcomes. This gives rise to so called public strategies. When the players use only public strategies then the set of possible equilibrium payoffs is same in any continuation game, i.e. after any history of play. Therefore, at any stage, the equilibrium payoffs can be decomposed into first-period payoffs and continuation payoffs which are again equilibrium payoffs. Abreu, Pearce, and Stachetti (1990) were first to apply dynamic programming principles to this idea and develop various characterizations of the equilibria payoffs. This work was further extended by Fudenberg, Levine, and Maskin (1994) and Fudenberg and Levine (1992) to obtain the folk theorems for their respective settings. But such a recursive structure is not available if monitoring is completely private and this is exactly what makes the private imperfect monitoring model much harder to analyse.

The second reason is the necessity of complex statistical inference. Namely, in order to determine what actions the opponent will take, any player i has to compute conditional probability distributions over the opponent's private histories given his own histories ($P(h^t_{-i}|h^t_i)$) which can be quite complicated for large t's.

8 Application to Online Settings

In this section we turn our attention to how to apply the discussed models to online settings. We start by motivating research on online reputation systems in general. Then we define the precise position of the game-theoretic modeling of reputation in the online world and outline some pros and cons of this kind of modeling. We conclude the section with a brief discussion on how to apply the game-theoretic models in P2P environments and what difficulties can arise.

8.1 Why Reputation Management?

The application to e-commerce markets offers probably the simplest possible answer to this question: in many online marketplaces reputation management is the only practically feasible way to encourage trustworthy behavior. Namely, if an online community is spanning the Globe then all classical assurance mechanisms, such as contractual agreements and litigation, are practically ineffective. This is particularly true if the the stakes involved are small in comparison with the costs incurred by these mechanisms. Then the only possible alternative is to track the users' behavior by eliciting feedback on their past doings from the community in hope that any misbehavior will be reported and thus disable or significantly decrease the future profit of the misbehavers.

eBay, the word's largest online auctioning site, offers some further insights. With respect to what opportunity might have been missed if no reputation management had been employed, it is sufficient to mention some recently announced figures from this site regarding the year 2003 (eBay (2004)): 971 million listings (from 94.9 million users in total), consolidated net revenues of $2.17 billion and the gross merchandise sales of $24 billion. On the other hand, some recent empirical studies (Resnick and Zeckhauser (2002), Houser and Wooders (2001) and Melnik and Alm (2002), just to mention a few) have shown that much of eBay's commercial success can be really attributed to its reputation mechanism (Feedback Forum) as a means of deterring dishonest behavior. To be exact, the analysis of eBay data carried out by Resnick and Zeckhauser (2002) has shown that "reputation profiles were predictive of future performance", while Houser and Wooders (2001) and Melnik and Alm (2002) performed a little bit more detailed analysis and came up with the conclusion that, as far as the sell of the goods the authors considered is concerned, Feedback Forum fulfilled its promises: the positive feedback of the sellers has been found to increase their price, while the negative one reduced it.

Without going into detail, we note that, though representative ones, e-commerce in general and eBay in particular are not the only examples of settings in need of the reputation management. There are many other settings in which at least some sort of reputation based social control mechanisms has been successfully employed. Examples range from Web search engines (Page, Brin, Motwani, and Winograd (1998)) and P2P networks to restaurant ratings.

8.2 Why and why not game-theoretic modeling?

eBay's Feedback Forum is just a well known example of what Resnick, Zeckhauser, Friedman, and Kuwabara (2000) call reputation systems and define as systems that help people decide whom to trust, encourage trustworthy behavior, and deter participation by those who are unskilled or dishonest through collecting, distributing, and aggregating feedback about the participants' past behavior.

A large number of online reputation mechanisms appeared in recent years. There are lots of web sites and P2P solutions allowing users to rate one another, leave feedback of any sort about other users, products and so on. However, we are not sure that a clear-cut systematic and analytic discipline of the reputation systems engineering can be identified here. This is exactly where game theory can help. Namely, common to the vast majority of the existing reputation systems is that the decision making process based on the feedback (either aggregated or raw) is intuitive, ad-hoc and error-prone. For example, how exactly can an eBay bidder use a seller profile to determine how much to bid and whether to bid at all? The power of game-theoretic reputation systems, as based on an analytic discipline, is precisely that they eliminate this problem: given that the players are rational utility maximizers then they cannot do anything better than playing exactly what is specified by an equilibrim, whichever is actually implemented. Further, with game-theoretic reputation models there is a clear link between the feedback aggregation strategy and the actual behavior of the participants, reflected in the chosen equilibrium. It is exactly this link what enables the mentioned systematism and analyticity. Put differently, by changing the feedback aggregation strategies reputation systems designers can tune the equilibria set according to their wishes and, in the ideal case, narrow down this set to only one equilibrium, presumably the socially most desirable one. Needless to say, this is exactly the main purpose of the reputation mechanisms.

With all this reasoning in mind we can derive quite precise steps in game-theoretically founded reputation mechanisms design:

- Analyzing interaction patterns in the considered community and identification of the stage game.
- Identification of the appropriate repeated game model (a long-run player against many short-run opponents or all players long-run or even a combination of these two).
- Defining of the feedback (past play) type along with a solution for incentivizing short-run players (if any) to leave it.
- Defining the way the feedback is aggregated.

In the rest of this section we will briefly describe how the current literature addresses the last two tasks. (We believe that the first two tasks are clear in themselves and we will not discuss them further.) As we saw, almost all discussed models involved short-run players sharing information about the long-run player's actions among themselves. In online settings, where there is a cost associated with communicating this information (leaving feedback), this cannot be taken for granted. Another difficulty is that even if the short-run players are incentivized to leave feedback how can we make sure that this is done truthfully. Miller, Resnick, and Zeckhauser (2002) analyse exactly this problem and propose a payment-based system based on the application of proper scoring rules to the reports. They prove that honest reporting is a Nash equilibrium and that, on the other hand, the budget is balanced by charging uninvolved players for payments that are based on the reports of others.

The last task in the above list is certainly the core one. Unfortunately, it turns out that in general it is not at all easy. Even deriving reputational equilibria can be quite complicated even when the stage game has a simple structure, the reverse task of devising a game with specific equilibria is in principle even more complicated. This is particularly true for private imperfect monitoring games and, unfortunately, exactly this class of games is the most frequent in online settings because any aggregation of the feedback (past play) makes the monitoring imperfect, while in the models with one long-run player playing against a sequence one-shot opponents (these models are of particular importance in huge online communities where repeated interactions between any two specific players are very improbable) it becomes necessarily private. This is exactly why the game-theoretic reputational models of online environments are very rare. Dellarocas (2003) presents one such model in which he models an eBay-like setting with a long-lived seller facing an infinite sequence of one-shot buyers. In each round the seller sells an item that can be of high and low quality and the buyers compete to buy the item in an auction. Because the seller decides which quality to deliver after receiving the payment the setting is with moral hazard. The author shows that if only a number of past outcomes is maintained and showed to the buyers it can be still the best strategy for the seller to deliver high quality and for the buyers to bid truthfully.

Apart from the mentioned complications that the game-theoretic modeling of reputation normally involves, there are also some very practical arguments against it. Thus, there is a concern about how well the described models approximate the real-world settings. For instance, is the model with one long-run player and many one shot opponents a good approximation of a typical eBay scenario in which one seller sells items to many buyers or this interaction cannot be considered in isolation from other interactions of the seller and the buyers? Some other concerns deal with the discounting criterion and the exact values of the players' discount factors. In the case we need them, how can we discover the discount factors or what happens if two or more (long-run) players have different factors. These practical issues are not usually studied in the literature.

Finally, there is a huge body of work proving that humans do not normally act as rational economic agents (Fehr and Gächter (2002), for example) raising the question of the practical usability of the game-theoretic modeling in general. More specifically, Bolton, Katok, and Ockenfels (2002) report on some peculiarities in how people treat reputation systems which can be again explained by their irrationality.

In P2P networks the hard problem we just described becomes even considerably harder. An implicit assumption of the above discussion was that the feedback aggregation was done by a central authority (eBay, for example) that did not have any incentive to distort it in any other way. Unfortunately, this assumption is not valid in P2P networks where the aggregation can be performed only by the peers themselves. But then the peers may find it profitable to distort the feedback or simply to reenter the network under a different identifier and abuse their knowledge of the complete feedback in the interactions it covers. To the best of our knowledge, there is no game-theoretic model that includes this possibility.

References

Abreu, D., D. Pearce, and E. Stachetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–1064.

Axelrod, R. (1984): *The Evolution of Cooperation*. Basic Books, New York.

Bolton, G., E. Katok, and A. Ockenfels (2002): "How effective are online reputation mechanisms? An experimental investigation," Discussion paper 25-2002, Max Planck Institute for Research into Economic Systems, Germany, ftp://ftp.mpiew-jena.mpg.de/pub/esi/discussionpapers/2002-25.pdf.

Border, K. C. (1985): *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge University Press, Cambridge, UK.

Cripps, M., G. Mailath, and L. Samuelson (2002): "Imperfect Monitoring and Impermanent Reputations," Working paper, Penn Institute for Economic Research.

Dellarocas, C. (2002): "The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Systems," Working paper, MIT.

——— (2003): "Efficiency and Robustness of Binary Feedback Mechanisms in Trading Environments with Moral Hazard," Working paper 4297-03, MIT.

eBay (2004): "eBay Inc. announces fourth quarter and full year 2003 financial results," http://investor.ebay.com/news/Q403/EBAY012104-712351.pdf.

Fehr, E., and S. Gächter (2002): "Altruistic punishment in humans," *Nature*, 14, 137–140.

Fudenberg, D., D. Kreps, and E. Maskin (1990): "Repeated Games with Long-run and Short-run Players," *Review of Economic Studies*, 57, 555–574.

Fudenberg, D., and D. Levine (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57(4), 759–778.

——— (1992): "Maintaining a Reputation when Strategies are Imperfectly Observed," *Review of Economic Studies*, 59, 561–579.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62(5), 997–1039.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or Incomplete Information," *Econometrica*, 54, 533–556.

Harsanyi, J. (1967–68): "Games with Incomplete Information Played by 'Bayesian' Players," *Management Science*, 14, 159–182, 320–334, 486–502.

Houser, D., and J. Wooders (2001): "Reputation in Auctions: Theory and Evidence from eBay," Working paper, University of Arizona.

Kandori, M. (2002): "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 102, 1–15.

Kreps, D., and R. Wilson (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253–279.

McAfee, R. P., and J. McMillan (1987): "Auctions and bidding," *Journal of Economic Literature*, 25, 699–738.

Melnik, M. I., and J. Alm (2002): "Does a Seller's Ecommerce Reputation Matter? Evidence from eBay Auctions," *Journal of Industrial Economics*, 50(3), 337–349.

Milgrom, P., and J. Roberts (1982): "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, 27, 280–312.

Miller, N., P. Resnick, and R. Zeckhauser (2002): "Eliciting Honest Feedback in Electronic Markets," Working paper, SITE02 workshop, Available at: http://www.si.umich.edu/ presnick/papers/elicit/.

Myerson, R. B. (1991): *Game Theory. Analysis of Conflict*. Harvard University Press.

Page, L., S. Brin, R. Motwani, and T. Winograd (1998): "The PageRank Citation Ranking: Bringing Order to the Web," Discussion paper, Stanford University, Stanford, CA.

Resnick, P., and R. Zeckhauser (2002): "Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system," in *The Economics of the Internet and E-Commerce*, ed. by M. R. Baye, vol. 11 of *Advances in Applied Microeconomics*. Amsterdam, Elsevier Science.

Resnick, P., R. Zeckhauser, E. Friedman, and K. Kuwabara (2000): "Reputation Systems," *Communications of the ACM*, 43(12), 45–48.

Selten, R. (1978): "The Chain-Store Paradox," *Theory and Decision*, 9, 127–159.