Learning to Play Network Games

Amy Greenwald

April 30, 1997

Contents

1 Introduction

In the classical theory of games, common knowledge of rationality, together with common knowledge of the underlying structure of a game, gives rise to the Nash equilibrium solution concept. If all the players are rational, and if they all know that they are all rational, and so on, then they all know that the others all play best responses, and as a result, they all play best responses to those best responses; this implies the Nash equilibrium solution concept. However, common knowledge of rationality and payoff functions are very strong assumptions. Throughout this paper, the common knowledge assumptions are dropped in varying capacities, thereby limiting the credibility of the Nash equilibrium solution concept in its usual form. This research focuses on alternative forms of equilibria which arise as a result of learning in repeated games in the absence of common knowledge.

Without the assumptions of common knowledge of rationality and payoff structure, games are not conducive to deductive solutions, such as Nash equilibrium. One of the focal points of modern game theory is inductive reasoning in repeated games, which is described as follows in Arthur [1]:

> Each agent keeps track of the performance of a private collection of belief-models. When it comes time to make choices, he acts upon his currently most credible (or possibly most profitable) one. The others he keeps at the back of his mind, so to speak. Alternatively, he may act upon a combination of several...Once actions are taken, agents update the track record of all their hypotheses.

This type of reasoning is known as belief-based learning. Examples of belief-baser learning algorithms include Bayesian updating and calibration. Under certain conditions, Bayesian learning converges to Nash equilibrium, while calibrated learning always converges to a generalization of Nash equilibrium known as correlated equilibrium. The intent of this thesis research is to develop efficient learning algorithms which quickly converge to reasonable approximations of equilibria in game-theoretic models of network routing and congestion problems.

2 Nash Equilibrium

If all the players in a game know the strategies and payoff functions of all the other players, and all the other players know that they know this information, and so on, and if all the players in a game are rational, and all the players know that all the players are rational, and so on, then the players deduce the optimal moves of the other players and their own optimal responses; this is precisely a Nash equilibrium. In this way, common knowledge of rationality and the underlying structure of a game gives rise to the traditional Nash equilibrium solution concept in strategic form games of complete information.

2.1 Examples

The most well-known game-theoretic scenario is the paradoxical situation known as the *Prisoner's Dilemma*, which was popularized by Axelrod [5] in his popular science book. The following is one variant of the story from which this game derives it name.[1]

An awful crime has been committed for which two suspects are being held incommunicado. The district attorney questions the two suspects. If both suspects confess, they are both punished, but not terribly severely, as the D.A. rewards them for their honesty (both payoffs equal 4). If only one suspect confesses, the confessor is severely punished for carrying out the crime singlehandedly (payoff equals 0), while the other suspect is let off scot free (payoff equals 5). Lastly, if neither suspect confesses, the D.A. has no choice but to convict both suspects, although the shared punishment is not as severe as the punishment of a single convict (both payoffs equal 1).

The Prisoner's Dilemma is a two player, strategic (or normal) form game. Such games are easily described by payoff matrices, where the strategies of player 1 and player 2 serve as column and row labels, respectively, and the corresponding payoffs are listed as pairs in matrix cells such that the first (second) number is the payoff to player 1 (2). A payoff matrix which describes the Prisoner's Dilemma is depicted in Figure 1, with C denoting "confess", or "cooperate", and D denoting "don't confess", or "defect".

[1] For the usual interpretation, see Rapoport [14], the two-time winner of the Prisoner's Dilemma computer tournament organized by Axelrod.

1 \ 2	C	D
C	4,4	0,5
D	5,0	1,1

Figure 1: The Prisoner's Dilemma

1 \ 2	H	T
H	1,-1	-1,1
T	-1,1	1,-1

Figure 2: Matching Pennies

This game is known as the Prisoner's Dilemma because the outcome of the game is (D, D), which yields suboptimal payoffs of $(1, 1)$. This equilibrium solution arises as a direct consequence of rationality. The reasoning is as follows. If player 1 plays C, then player 2 is better off playing D, since D yields a payoff of 5, whereas C yields only 4; but if player 1 plays D, then player 2 is again better off playing D, since D yields a payoff of 1, whereas C yields 0. Hence, regardless of the strategy of player 1, a rational player 2 plays D. By a symmetric argument, a rational player 1 also plays D. Thus, the outcome of the game is (D, D).

A second well-known example of a two-player game is a game called *Matching Pennies*. In this game, each of the two players flips a coin, and the payoffs are determined as follows (see Figure 2). Let player 1 be the *matcher*, and let player 2 be the *mismatcher*. If the coins come up matching (*i.e.*, both heads or both tails), then player 2 pays player 1 the sum of \$1. Otherwise, if the coins do not match (*i.e.*, one head and one tail), then player 1 pays player 2 the sum of \$1. This is an example of a zero-sum game where the interests of the players are diametrically opposed; this class of games is so-called because the payoffs in the matrix do indeed sum to zero.

1 \ 2	*B*	*F*
B	2,1	0,0
F	0,0	1,2

Figure 3: Battle of the Sexes

Another popular two-player game is called the *Battle of the Sexes.* A man and a woman would like to spend an evening out together; however, the man prefers to go to a football game (strategy F), while the woman prefers to go to the ballet (strategy B). Both the man and the woman prefer to be together, even at the event that is not to their liking, rather than go out alone. The payoffs of this coordination game are shown in Figure 3; the woman is player 1 and the man is player 2.

The last example that is presented is an ecological game which was studied by Maynard Smith [15] in his analysis of the theory of evolution in terms of games. The game is played between animals of similar physique who live in the wilderness and encounter one another in their search for prey. During an encounter between two animals, each animal has a choice between behaving as a hawk: *i.e.*, fighting for the prey; or as a dove: *i.e.*, running away peacefully. If both animals decide to play like hawks, then each animal has an equal chance of winning the value v of the prey or of losing the fight at cost c, where $0 < v < c$; thus, the expected payoff to both players is $(v - c)/2$. Alternatively, if both animals act as doves, then the prey is shared with equal payoffs $v/2$. Finally, if one animal behaves like a hawk and the other behaves like a dove, then the hawk gets a payoff worth the full value of the prey and the other gets nothing. In this game, the animals prefer to choose opposing strategies: if one animal plays hawk, then it is in the best interest of the other to play dove; and inversely, if one animal plays dove, then it is in the best interest of the other to play hawk.

This section included several popular examples of strategic form games. The next section presents the formal theory of strategic form games and the Nash equilibrium solution concept, which follows directly from common knowledge.

1 \ 2	H	D
H	(v-c)/2, (v-c)/2	v,0
D	0,v	v/2,v/2

Figure 4: Hawks and Doves

2.2 Strategic Form Games

This section develops the formal theory of finite games in strategic form. Let $\mathcal{I} = \{1, \ldots, I\}$ be a set of *players*, where $I \in \mathbb{N}$ is the number of players. The (finite) set of *pure strategies* available to player $i \in \mathcal{I}$ is denoted by S_i, and the set of pure strategy profiles is the cartesian product $S = \prod_i S_i$. By convention, write $s_i \in S_i$ and $s = (s_1, \ldots, s_I) \in S$. In addition, let $S_{-i} = \prod_{j \neq i} S_j$ with element $s_{-i} \in S_{-i}$, and write $s = (s_i, s_{-i}) \in S$. The payoff (or reward) function $r_i : S \to \mathbb{R}$ for the i^{th} player is a real-valued function on S; in this way, the payoffs to player i depend on the strategic choices of all players. This description is summarized in the following definition.

Definition 2.1 *A* strategic form game Γ *is a tuple*

$$\Gamma = (\mathcal{I}, (S_i, r_i)_{i \in \mathcal{I}})$$

where

- $\mathcal{I} = \{1, \ldots, I\}$ *is a set of players* $(i \in \mathcal{I})$
- S_i *is a finite strategy set* $(s_i \in S_i)$
- $r_i : S \to \mathbb{R}$ *is a payoff function*

Example 2.1 *Formally, the Prisoner's Dilemma consists of a set of players* $\mathcal{I} = \{1, 2\}$, *with strategy sets* $S_1 = S_2 = \{C, D\}$, *and payoffs as follows:* $r_1(C, C) = r_2(C, C) = 4$, $r_1(C, D) = r_2(D, C) = 0$, $r_1(D, C) = r_2(C, D) = 5$, *and* $r_1(D, D) = r_2(D, D) = 1$.

A Nash equilibrium is a strategy profile from which none of the players has any incentive to deviate. In particular, no player can achieve strictly greater payoffs by choosing any strategy other than the one prescribed by the profile, given that all other players choose their prescribed strategies. In this sense, a Nash equilibrium specifies optimal strategic choices for all players.

In the Prisoner's Dilemma, (D, D) is a Nash equilibrium because given that player 1 plays D, the best response of player 2 is to play D; and, given that player 2 plays D, the best response of player 1 is to play D. The Battle of the Sexes has two pure strategy Nash equilibria, namely (B, B), and (F, F), by the following reasoning. If the woman chooses B, then the best response of the man is B; and if the man chooses B, then the best response of the woman is B. Analogously, if the woman chooses F, then the best response of the man is F; and if the man chooses F, then the best response of the woman is F. Note that the game of Matching Pennies does not have a pure strategy Nash equilibrium. If player 1 plays H, then the best response of player 2 is T; but if player 2 plays T, the best response of player 1 is not H, but T. Moreover, if player 1 plays T, then the best response of player 2 is H; but if player 2 plays H, then the best response of player 1 is not T, but H. However, this game does have a mixed strategy Nash equilibrium. A mixed strategy is a randomization over a set of pure strategies. In particular, the probabilistic strategy profile in which both players choose H with probability $\frac{1}{2}$ and T with probability $\frac{1}{2}$ is a mixed strategy Nash equilibrium.

A mixed strategy set for player i (notation Q_i) is the set of probability distributions over the pure strategy set S_i: *i.e.*,

$$Q_i = \{q_i : S_i \to [0,1] \mid \sum_{s_i \in S_i} q_i(s_i) = 1\}$$

The notational conventions extend to mixed strategies, *e.g.*, $q = (q_i, q_{-i}) \in Q$. In the context of mixed strategies, the expected payoffs to player i, from strategy profile q, is given by:

$$E[r_i(q)] = \sum_{s_i \in S_i} q_i(s_i) \cdot r_i(s_i, q_{-i})$$

As usual, the payoffs to player i depend on the mixed strategies of all players.

Definition 2.2 *Given an opposing strategy $q_{-i} \in Q_{-i}$. A strategy $q_i^* \in Q_i$ is* rational *for player i iff for all strategies $q_i \in Q_i$,*

$$E[r_i(q_i^*, q_{-i})] \geq E[r_i(q_i, q_{-i})]$$

Definition 2.3 *A strategy profile $q^* = (q_i^*, q_{-i}^*)$ is a* Nash equilibrium *iff strategy q_i^* is rational, given opposing strategy q_{-i}^*, for all players $i \in \mathcal{I}$.*

An implication of the assumption of rationality is that a rational player always plays a *best response* to the strategies of the other players, where a best response is an optimizing strategy. A Nash equilibrium is a strategy profile in which all players choose strategies that are best responses to the strategic choices of the other players. An alternative characterization of Nash equilibrium is given in terms of best response sets.

Definition 2.4 *The set of best responses for player i to strategy profile q is:*

$$\mathrm{BR}_i(q) = \{q_i^* \in Q_i \mid E[r_i(q_i^*, q_{-i})] \geq E[r_i(q_i, q_{-i})], \forall q_i \in Q_i\}$$

Let $\mathrm{BR}(q) = \prod_i \mathrm{BR}_i(q)$.

A Nash equilibrium is a strategy profile in which the players all play best responses to all the other players' strategies.

Definition 2.5 *A Nash equilibrium is a strategy profile q^* s.t. $q^* \in \mathrm{BR}(q^*)$.*

It is apparent from this definition that a Nash equilibrium is a fixed point of the best response relation. The proof of existence of Nash equilibrium utilizes a fundamental result in topology: namely, Kakutani's fixed point theorem, which is a generalization of Brouwer's fixed point theorem.

Theorem 2.1 (Nash, 1951) *Every finite, strategic form game has a mixed strategy Nash Equilibrium.*

Although Nash equilibrium is the generally accepted solution concept in the deductive analysis of strategic form games, it should be noted that the

Nash equilibrium solutions in the stated examples are somewhat peculiar. In particular, in the Prisoner's Dilemma, the Nash equilibrium payoffs are sub-optimal. Moreover, in the game of Matching Pennies, the Nash equilibrium solution is probabilistic. Finally, in the Battle of the Sexes and the game of Hawks and Doves, the Nash equilibrium is not unique. In view of these quirky outcomes, this thesis considers alternative forms of equilibria which arise as a result of various learning processes in repeated games.

3 Other Equilibria

In this section, two generalizations of the Nash equilibrium solution concept are introduced and their relationship is explored. Bayesian Nash equilibrium arises in games of incomplete information when players maximize expected payoffs with respect to beliefs. Correlated equilibrium generalizes Nash by allowing for possible dependencies in strategic choices. Information games provide an appropriate framework in which to interpret these equilibria.

3.1 Information games

An information game is a strategic form game in which players maintain a database of knowledge and beliefs about the state of the world. This information is stored in a knowledge belief system, which is defined as follows.

Definition 3.1 *A* knowledge belief system *is a probability space*

$$\mathcal{K} = (\Omega, (\mathcal{P}_i, \pi_i)_{i \in \mathcal{I}})$$

where

- Ω *is a finite set of possible states of the world* ($\omega \in \Omega$)
- $\mathcal{P}_i$ *is an information partition*[2] *on* Ω ($P_i \in \mathcal{P}_i$)
- p_i *is a prior probability*[3] *on* Ω

[2] Technically, $\mathcal{P}_i$ is a σ-field.

[3] Assume p_i is measurable: *i.e.*, for all $i \in \mathcal{I}$, for all $\omega, \omega' \in \Omega$, $\pi_i(\omega) = \pi_i(\omega')$ whenever $P_i(\omega) = P_i(\omega')$.

An element of an information partition $\mathcal{P}_i$ is called an information set of player i at state ω, and is denoted by $P_i(\omega)$. Intuitively, $P_i(\omega)$ is an equivalence class consisting of those states that are indistinguishable from ω from the point of view of player i.

Example 3.1 *Consider the state of knowledge today about the price of* IBM *stock tomorrow for two players. In this scenario, the possible states of the world are up and down:* i.e., $\Omega = \{U, D\}$*. The information partition of both players is the trivial partition, namely* $\{\emptyset, \Omega\}$*, since neither player knows the state of the world tomorrow. However, the prior probabilities, or beliefs, of the two players need not agree. For example, player* 1 *may attribute equal prior probabilities to both up and down:* i.e., $\pi_1(U) = \pi_2(D) = \frac{1}{2}$*; while player* 2 *might attribute prior probabilities of* $\frac{1}{3}$ *to up and* $\frac{2}{3}$ *to down:* i.e., $\pi_2(U) = \frac{1}{3}$ *and* $\pi_2(D) = \frac{2}{3}$.

Definition 3.2 *An* information game $\Gamma_{\mathcal{K}}$ *is a strategic form game* Γ *together with a knowledge belief system* $\mathcal{K}$.

Example 3.2 Consider the Battle of the Sexes, which is described in the formal framework as follows. The set of players $\mathcal{I} = \{W, M\}$, with strategy sets $S_W = S_M = \{B, F\}$, and payoffs as follows:

$$\begin{array}{ll} r_W(B,B) = r_M(F,F) = 2 & r_W(F,F) = r_M(B,B) = 1 \\ r_W(B,F) = r_M(B,F) = 0 & r_W(F,B) = r_M(F,B) = 0 \end{array}$$

In the Battle of the Sexes viewed as an information game, the set of states of the world consists of all possible outcomes of the strategic form game: *i.e.*, $\Omega = \{(B,B),(B,F),(F,B),(F,F)\}$. If the woman is playing strategy B, then the woman's knowledge of the world is described by information partition $\mathcal{P}_W = \{\emptyset, \Omega, \{(B,B),(B,F)\}, \{(F,B),(F,F)\}\}$. In addition, if the man is playing strategy B, then the man's knowledge of the world is described by information partition $\mathcal{P}_M = \{\emptyset, \Omega, \{(B,B),(F,B)\}, \{(B,F),(F,F)\}\}$. In this case, the woman will attribute prior probabilities p_W to state (B,B) and $1 - p_W$ to state (B,F), according to her beliefs about the man's strategy, while the man will attribute prior probabilities p_M to state (B,B) and $1-p_M$ to state (F,B), according to his beliefs about the woman's strategy.

1 \ 2	B	F
B	1/2	0
F	0	1/2

Figure 5: The Battle of the Sexes Revisited

3.2 Correlated Equilibrium

The notion of correlated equilibrium generalizes that of Nash equilibrium, by eliminating the independence condition among the mixed strategies of the players. In particular, correlated equilibrium allows for dependencies in the players' choices of randomizations over their pure strategies. A daily example of a correlated equilibrium is a traffic light; red (green) traffic signal suggests that cars should stop (go), and in fact, these suggestions are best responses to the simultaneous suggestions for the actions of others.

As another example, consider once again the Battle of the Sexes. This game has three Nash equilibria, two of which are pure strategy equilibria, namely (B, B) and (F, F), and the mixed strategy $(\frac{2}{3}, \frac{1}{3})$ for the woman and $(\frac{1}{3}, \frac{2}{3})$ for the man, which yields equal expected payoffs of $(\frac{2}{3}, \frac{2}{3})$ to both. A correlated equilibrium of this game which yields expected payoffs of $(1\frac{1}{2}, 1\frac{1}{2})$ is given by the distribution $(\frac{1}{2}(B, B), \frac{1}{2}(F, F))$. Figure 5 presents this correlated equilibrium in convenient form.

In general, it is possible to achieve correlated equilibrium payoffs from any convex combination of Nash equilibrium. Moreover, it is also possible to achieve payoffs via correlated equilibrium outside the convex hull of Nash equilibrium payoffs. For example, in the game depicted in Figure 6, the Nash equilibrium payoffs achieved via mixed strategies $(\frac{1}{5}, \frac{4}{5})$ and $(\frac{4}{5}, \frac{1}{5})$ for players 1 and 2, respectively, yield expected payoffs of 4 for both players. In contrast, the correlated equilibrium strategies presented in Figure 6 generate expected payoffs of $4\frac{1}{4}$.

1 \ 2	L	R
T	4,4 1/2	4,5 1/4
B	5,4 1/4	0,0 0

Figure 6: Correlated Equilibrium

In information games, expected payoffs are computed in terms of prior probabilities, or beliefs. Given pure strategy profile s, the payoff that player j expects to be rewarded to player i is as follows:

$$E_j[r_i(s)] = p_j(s) \cdot r_i(s)$$

A strategy is rational for player i which maximizes i's expectation of i's payoffs. A player is rational who plays only rational strategies.

Definition 3.3 *Given an opposing strategy $s_{-i} \in S_{-i}$. A strategy $s_i^* \in S_i$ is* rational *for player i iff for all strategies $s_i \in S_i$,*

$$E_i[r_i(s_i^*, s_{-i})] \geq E_i[r_i(s_i, s_{-i})]$$

The following two definitions make explicit the relationship between Nash equilibrium and correlated equilibrium, by defining correlated equilibrium and redefining Nash equilibrium, both in terms of information games.

Definition 3.4 *Given an information game. A strategy profile $s^* = (s_i^*, s_{-i}^*)$ is a* correlated equilibrium *iff for all players $i \in \mathcal{I}$, strategy s_i^* is rational, given opposing strategy s_{-i}^*, and the common prior assumption holds.*

Definition 3.5 *Given an information game. A strategy profile $s^* = (s_i^*, s_{-i}^*)$ is a* Nash equilibrium *iff for all players $i \in \mathcal{I}$, strategy s_i^* is rational, given opposing strategy s_{-i}^*, the common prior assumption holds:* (i.e., $p_i = p$)*, and the information partitions $\mathcal{P}_i$ are independent.*

3.3 Bayesian Nash Equilibrium

This section considers Bayesian games of incomplete information, in which the missing piece of information is the common payoff structure of the game. A well-known game of incomplete information is the envelope paradox. A father offers each of his two sons an envelope with either $\$10^m$ or $\$10^n$, where $|m - n| = 1$, for $0 \leq m, n \leq 6$. Each brother can accept his envelope, or chose to engage in a bet in which he pays the father \$1 for the right to swap envelopes with his brother, provided that his brother has also chosen to engage in this bet. Otherwise, he simply loses \$1. In this scenario, the payoffs to the brothers are private; thus, the envelope paradox can be modeled as a Bayesian game. The missing information regarding payoffs is described in terms of an unknown state of the world. In particular, let (m, n) denote the state of the world, where m (n) is the exponent of the payoff to the first (second) brother. The pure strategy sets of the brothers are BET and NO BET. The payoff matrix in Figure 7 depicts the outcomes of the envelope game in terms of the unknown state of the world. The envelope paradox is so-called because regardless of what the father gives to his sons, it appears that it is always in both of their best interests to accept the bet. In general, since the probablity that the second brother receives 10 times as much money as the first brother is $\frac{1}{2}$, the expected value of the second brother's lot, given the information that the first brother has about his own lot, is always approximately 5 times greater than the first brother's lot. Thus, it is in the first brother's best interest to bet. The reasoning is analogous for the second brother. This paradox is resolved via game-theoretic reasoning.

Definition 3.6 *A* Bayesian game *is an information game in which payoffs are a function of the state of the world. In particular, $r_i : \Omega \times S \to \mathbb{R}$, for all players $i \in \mathcal{I}$.*

Bayesian rationality is an extension of the usual economic understanding of rationality in which players choose strategies that maximize their expected payoffs, according to their beliefs. Given strategy profile s, the expected payoffs to player i at state ω_0 as predicted by player j is based on player j's beliefs, which are described j's information partition $\mathcal{P}_j$:

$$E_j[r_i(s, \omega_0)|\mathcal{P}_j] = \sum_{\omega \in P_j(\omega_0)} p_j(\omega|P_j(\omega_0)) \cdot r_i(s, \omega)$$

1 \ 2	B	NB
B	10^n-1, 10^m-1	10^m-1, 10^n
NB	10^m, 10^n-1	10^m, 10^n

Figure 2

Figure 7: The Envelope Paradox

A strategy is Bayesian rational for player i which maximizes i's expectation of i's payoffs, given i's beliefs. A player is Bayesian rational who plays only Bayesian rational strategies.

Definition 3.7 *Given an opposing strategy $s_{-i} \in S_{-i}$. A strategy $s_i^* \in S_i$ is* Bayesian rational *for player i iff given information partition $\mathcal{P}_i$, for all $\omega \in \Omega$, for all strategies $s_i \in S_i$,*

$$E_i[r_i((s_i^*, s_{-i}), \omega)|\mathcal{P}_i] \geq E_i[r_i((s_i, s_{-i}), \omega)|\mathcal{P}_i]$$

Definition 3.8 *Given a Bayesian game. A strategy profile $s^* = (s_i^*, s_{-i}^*)$ is said to be a* Bayesian Nash equilibrium *iff strategy s_i^* is Bayesian rational, given opposing strategy s_{-i}^*, for all players $i \in \mathcal{I}$, and $p_i = p$, for all $i \in \mathcal{I}$.*

The strategy profile (NB, NB) is the unique Bayesian Nash equilibrium in the Bayesian game which depicts the envelope paradox. The game is such that the brother's have common prior beliefs about the possible states of the world. Now, given that the first brother chooses not to bet, the second brother also chooses not to bet, since he incurs a loss of \$1 otherwise; this is the case at all states of the world. The situation is symmetric for the second brother. Moreover, these strategic choices form the unique Bayesian Nash equilibrium, since the strategy (B, B) is necessarily a strictly sub-optimal choice for one of the brothers.

4 Belief-Based Models

In strategic form games of complete information, deductive reasoning gives rise to Nash equilibrium, assuming common knowledge of rationality and the underlying payoff structure. However, the common knowledge assumption is very strong and often unrealistic in real-world games, such as internet and other network games. Consequently, this assumption is dropped in varying capacities throughout the remainder of this paper. In order to reconstruct a small part of this lost information, the situations considered allow players to learn via repeated play of games. This section considers two examples of belief-based learning models, namely Bayesian updating and calibrated learning.

4.1 Repeated Games

This section describes belief-based models of learning in repeated games (see Figure 8). Players have an initial set of beliefs. Given these beliefs, players *play* the game: *i.e.*, players choose a strategy from their respective strategy sets. Then, based on the actual strategies that are played, players *learn*: *i.e.*, update their beliefs. Recall that a mixed strategy $q_i \in Q_i$ for the ith player is a probability distribution over the set of pure strategies. On the other hand, a belief is a probability distribution over the set of opponents' pure strategies, denoted $p_{-i} \in Q_{-i}$. Intuitively, a strategy implicitly encodes how players behave as they learn from opponents' past actions. Likewise, a belief records how players think other players will behave as the other players learn about the repeated play of the game.

Definition 4.1 *A* history *of length t is a sequence of plays drawn from S: $h^t = (s^1, \ldots, s^t)$. Let H^t denote the set of all t-histories, and let $H = \bigcup_t H^t$.*

The play of the game is prescribed by behavioral strategies, and the way in which players learn is given by a learning rule. In repeated games, a behavioral strategy for the ith player is a function $g_i : H \to Q_i$ from the set of all possible histories to the ith mixed strategy set. Likewise, a learning rule is for the ith player is a function $f_i : H \to Q_{-i}$ from the set of all possible histories to the set of opponents' mixed strategy sets.

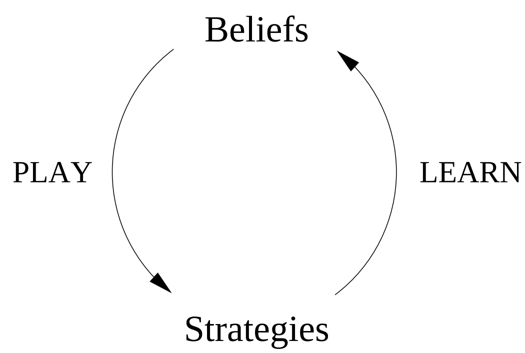

Figure 8: Belief-Based Model

Definition 4.2 *A* learning process *is a family of pairs* $\{(f_i, g_i)_{i\in\mathcal{I}}\}$ *where* f_i *is a learning rule and* g_i *is a behavioral strategy for the* i^{th} *player.*

Definition 4.3 *A learning process induces a* learning path*, which is defined as a sequence* $\{(h^t, p^t_{-i}, q^t_i)_{i\in\mathcal{I}} \mid t = 0, 1, \ldots\}$*, where*

- h^t *is a t-history (assume* $h^0 = \emptyset$*)*
- $p^t_{-i} = f_i(h^{t-1})$ *is player i's belief at time t*
- $q^t_i = g_i(h^{t-1})$ *is player i's strategy at time t*

The remainder of this paper focuses on different learning processes and the corresponding solution concepts which arise as a result of repeated play of information games, in the absence of common knowledge. The rationality assumption is maintained throughout this section in the following formalism:

$$\forall i, \forall t,\ q^t_i \in \ \mathrm{BR}_i(p^t_{-i})$$

In other words, all players choose best responses to their beliefs, at all times. More specifically, given player i's beliefs p^t_{-i} at time t, the best response set for player i is:

$$\begin{aligned}
& \mathrm{BR}_i(p^t_{-i}) \\
= \ & \{q^*_i \in Q_i \mid E_i[r_i(q^*_i, q_{-i})] \geq E_i[r_i(q_i, q_{-i})], \forall q_i \in Q_i\} \\
= \ & \{q^*_i \in Q_i \mid p^t_{-i}(q^*_i, q_{-i}) \cdot r_i(q^*_i, q_{-i}) \geq p^t_{-i}(q_i, q_{-i}) \cdot r_i(q_i, q_{-i}), \forall q_i \in Q_i\} \\
= \ & \{q^*_i \in Q_i \mid \textstyle\sum_{h^t_{q^*}} p^t_{-i}(h^t) \cdot r_i(h^t) \geq \sum_{h^t_q} p^t_{-i}(h^t) \cdot r_i(h^t), \forall q_i \in Q_i\}
\end{aligned}$$

where $\sum_{h^t_q}$ is the sum over all histories induced by strategy profile q.

4.2 Bayesian Learning

This section discusses convergence to Nash equilibrium. The principal result is that learning processes which satisfy an asymptotic property known as merging converge to Nash equilibrium. Moreover, Bayesian learning indeed satisfies the merging property. In the following section, a generalization of merging known as calibration is considered, and convergence to correlated equilibrium, a generalization of Nash equilibrium, is established.

Lemma 4.1 *A learning path* $\{(h^t, p^t_{-i}, q^t_i)_{i \in \mathcal{I}} \mid t = 0, 1, \ldots\}$ converges to Nash equilibrium *if the following mathematical conditions are satisfied:*

- RATIONALITY*:* $\forall i, \forall t, \; q^t_i \in \text{BR}_i(p^t_{-i})$
- COMMON PRIOR ASSUMPTION*:* $\forall i \neq j, \lim_{t \to \infty} \; || \; p^t_{-i} - p^t_{-j} \; || = 0$
- INDEPENDENCE*:* $\forall i \neq j$*,* $\mathcal{P}_i$ and $\mathcal{P}_j$ are pairwise independent

A learning rule satisfies the merging property if beliefs eventually agree with, or merge into, the truth. In particular, if for all players i and j, the probability measures $(p_{-i})_j$ and q_j eventually coincide, then the merging property holds for the corresponding learning process.

Definition 4.4 *A learning path* $\{(h^t, p^t_{-i}, q^t_i)_{i \in \mathcal{I}} \mid t = 0, 1, \ldots\}$ *satisfies the* merging property *iff* $\forall i \neq j$*,*

$$\lim_{t \to \infty} \; || \; (p^t_{-i})_j - q^t_j \; || = 0$$

The following theorem states that if players are rational, and if a learning path in a repeated game satisfies the merging property, then the learning path converges to Nash equilibrium.

Theorem 4.1 (Kalai and Lehrer, 1990) *Given a game* $\Gamma_{\mathcal{K}}$ *that is played repeatedly among rational players. If a learning process* $\{(f_i, g_i)_{i \in \mathcal{I}}\}$ *induces a learning path* $\{(h^t, p^t_{-i}, q^{t*}_i)_{i \in \mathcal{I}} \mid t = 0, 1, \ldots\}$ *for which the merging property holds, then* $q^{t*} = (q^{t*}_i)_{i \in \mathcal{I}}$ *is a Nash equilibrium.*

Proof 4.1 (Idea) By assumption, all players are rational. Thus, it suffices to show that merging implies independence in information partitions and the common prior assumption. The merging property holds when players learn the truth about other players' mixed strategies. At that point, players' strategic choices do not affect their beliefs about other players' distributions: *i.e.*, the knowledge among the players is independent.

Common priors follow from the merging property via the following:

$$\begin{array}{ll} \forall i \neq j \neq k, & \lim_{t\to\infty} \; || \; (p^t_{-i})_k - q^t_k \; || = 0 \\ and & \lim_{t\to\infty} \; || \; (p^t_{-j})_k - q^t_k \; || = 0 \\ \Rightarrow & \lim_{t\to\infty} \; || \; (p^t_{-i})_k - (p^t_{-j})_k \; || = 0 \end{array}$$

The lemma below states that if initial beliefs satisfy a grain of truth, then learning via Bayesian updating satisfies the merging property. A player's beliefs are said to contain a grain of truth if they assign positive probability to the actual strategies employed by the other players. It follows from this lemma together with the theorem above that a learning path that is induced by a Bayesian learning rule asymptotically approaches a Nash equilibrium path of play.

Definition 4.5 *Given time t and strategy profile q^t_{-i} for players other than player i. Player i's beliefs, namely p^t_{-i}, are said to* contain a grain of truth *iff* $\exists \hat{q}^t_{-i} \in Q_{-i}, 0 < \alpha \leq 1$ s.t.

$$p^t_{-i} = \alpha q^t_{-i} + (1-\alpha)\hat{q}^t_{-i}$$

Lemma 4.2 (Kalai and Lehrer, 1990) *If for all players $i \in \mathcal{I}$, the initial beliefs $p^0_{-i} \in Q_{-i}$ contain a grain of truth and the learning rule f_i is Bayesian updating, then the learning path $\{(h^t, p^t_{-i}, q^t_i)_{i\in\mathcal{I}} \mid t = 0, 1, \ldots\}$ induced by the learning process $\{(f_i, g_i)_{i\in\mathcal{I}}\}$ satisfies the merging property.*

Proof 4.2 (Idea) The sequence of posterior probabilities that result from Bayesian updating satisfies a property like consistent estimation, which says that this sequence tends to get closer and closer to the actual distributions of play, as t increases, provided that actual strategies are in the support of players' prior belief distributions. In other words, the belief distributions eventually merge with the actual distributions of play.

The above results show that repeated play of information games among rational players converges to approximate Nash equilibrium provided that players' initial beliefs assign positive probability to the strategies that their opponents actually play: *i.e.*, contain a grain of truth.

4.3 Calibrated Learning

The previous section considered a merging property which arises as a result of Bayesian learning in repeated information games. This section considers a statistical notion known as calibration, which is used to gauge the credibility of forecasting methods, such as weather predictors. Recall that a forecast satisfies the merging property if forecast beliefs converge to actual empirical frequencies. A calibrated forecast is one in which beliefs converge to empirical frequencies which are *conditioned on those beliefs*. Calibration is weaker than merging; thus, calibrated learning processes converge to a more general equilibrium concept than Nash, namely correlated equilibrium.

Dawid [7] gives the following intuitive description of calibration:

> Suppose that, in a long sequence of weather forecasts, we look at all those days for which the forecast probability of precipitation was, say, close to some given value p and (assuming these form an infinite sequence) determine the long run proportion ρ of such days on which the forecast event (rain) in fact occurred. The plot of ρ against p is termed the forecaster's *empirical calibration curve.* If the curve is the diagonal $\rho = p$, the forecaster may be termed *well-calibrated.*

For example, suppose it rains every other day; the empirical distribution of rain is given by the sequence $1, 0, 1, 0, \ldots$. The forecast $1, 0, 1, 0, \ldots$ is well-calibrated, since it rains with probability 1 on the days in which rain is predicted with probability 1. On the other hand, a daily forecast of rain with probability $\frac{1}{2}$ is also well-calibrated, since it rains on exactly half the days in which the forecast probability equals $\frac{1}{2}$. The latter example demonstrates that calibration is an extremely weak requirement; as such, it is often viewed as a minimal necessary condition of reliable forecasting techniques.

Game-theoretically, beliefs are well-calibrated with empirical frequencies iff for all histories h, for all $0 \leq p \leq 1$,

$$E[q_{-i}(h|p_{-i}(h) = p)] = p$$

In other words, the expected empirical probability q_{-i} of history h, given that beliefs p_{-i} attribute probability p to history h, is in fact equal to p. Let $P[(q^t|p^t_{-i})_j]$ denote the conditional probability that the strategy q^t_j is the actual randomization that is played by player j at time t, given beliefs player i's beliefs about j, namely $(p^t_{-i})_j$. In particular, if δ is the indicator function,

$$P[(q^t|p^t_{-i})_j] = \frac{\#((p^t_{-i})_j = p) \cdot \delta_{q^t_j, p}}{\#((p^t_{-i})_j = p)}$$

Definition 4.6 *A learning path* $\{(h^t, p^t_{-i}, q^t_i)_{i \in \mathcal{I}} \mid t = 0, 1, \ldots\}$ *is* calibrated *iff* $\forall i \neq j$,

$$\lim_{t \to \infty} \; || \; (p^t_{-i})_j - P[(q^t|p^t_{-i})_j] \; || = 0$$

It follows from this definition that merging implies calibration, since true probability is stronger than conditional probability. Analogous to the result that a learning process which satisfies the merging property converges to Nash equilibrium, a calibrated learning process, gives rise to a correlated equilibrium.

Lemma 4.3 *A learning path* $\{(h^t, p^t_{-i}, q^t_i)_{i \in \mathcal{I}} \mid t = 0, 1, \ldots\}$ converges to correlated equilibrium *if the following two conditions hold:*

- RATIONALITY*:* $\forall i, \forall t, \; q^t_i \in \mathrm{BR}_i(p^t_{-i})$
- COMMON PRIOR ASSUMPTION*:* $\forall i \neq j, \lim_{t \to \infty} \; || \; p^t_{-i} - p^t_{-j} \; || = 0$

Theorem 4.2 (Foster and Vohra, 1995) *Given a game* $\Gamma_{\mathcal{K}}$ *that is played repeatedly among rational players. If a learning process* $\{(f_i, g_i)_{i \in \mathcal{I}}\}$ *induces a calibrated learning path* $\{(h^t, p^t_{-i}, q^{t*}_i)_{i \in \mathcal{I}} \mid t = 0, 1, \ldots\}$*, then* $q^{t*} = (q^{t*}_i)_{i \in \mathcal{I}}$ *is a correlated equilibrium.*

Proof 4.3 (Idea) Calibrated learning implies that beliefs eventually agree with empirical frequencies. But empirical frequencies necessarily agree; thus, beliefs eventually agree, satisfying the common prior assumption. Now by assumption, all players are rational. Therefore, since players are rational and have common prior beliefs, it follows that calibrated learning converges to correlated equilibrium.

More formally, calibrated learning implies the common prior assumption via the following reasoning. The crucial observation is that the expected long-run frequencies, conditioned on beliefs, are the actual long-run frequencies, and therefore must agree. Moreover, these frequencies yield a correlated equilibrium.

$$
\begin{array}{ll}
\forall i \neq j \neq k, & \lim_{t\to\infty} \ || \ (p^t_{-i})_k - P(q|p_{-i})^t_k \ || = 0 \\
\mathit{and} & \lim_{t\to\infty} \ || \ (p^t_{-j})_k - P(q|p_{-j})^t_k \ || = 0 \\
\Rightarrow & \lim_{t\to\infty} \ || \ (p^t_{-i})_k - q^{t*}_k \ || = 0 \\
\mathit{and} & \lim_{t\to\infty} \ || \ (p^t_{-j})_k - q^{t*}_k \ || = 0 \\
\Rightarrow & \lim_{t\to\infty} \ || \ (p^t_{-i})_k - (p^t_{-j})_k \ || = 0
\end{array}
$$

This section described several formal results regarding belief-based models of learning in repeated information games. In particular, learning processes which satisfy the merging property necessarily converge to Nash equilibrium, while calibrated learning processes give rise to correlated equilibrium. One of the primary goals of this research is to exhibit efficient algorithms that satisfy properties such as merging and calibration in order to guarantee convergence to equilibrium solutions in network games, such as those presented in the final section.

5 Network Games

This section presents three examples of interesting network games. The first is the Santa Fe bar game, which is an abstraction of the problem of routing packets over a network. This game is analyzed without assuming common knowledge of rationality. Secondly, a congestion game is introduced that serve as a more realistic model of network interaction. The congestion game assumes neither common knowledge of rationality nor of payoffs; moreover, players do not even know their own payoffs. The last game is a traditional problem in economics, known as the free-rider problem, which is relevant in a network setting. This problem is described as a Bayesian game of incomplete information in which payoffs (costs) are private.

5.1 Santa Fe Bar Problem

This section describes an interesting example of a repeated game, namely the Santa Fe bar problem. This game is one of complete information that affords a deductive solution, assuming common knowledge of rationality. In addition, computer simulations of a form of inductive reasoning give rise to convergent behavior, in the absence of common knowledge, but in the presence of a small random component.

The *Santa Fe bar problem* was invented by Brian Arthur [1], an economist at the Santa Fe Institute, in order to illustrate how one might model inductive reasoning. Here is the scenario, in a nutshell:

> There is one bar in Santa Fe. Every night, the city dwellers make a decision as to whether or not they will go to the bar. They would all like to be at the bar if and only if it is not too crowded. More specifically, if there is an available seat, they all prefer to go to the bar, but if not, they all prefer to stay at home.

The Santa Fe bar problem is a non-cooperative game. As such, it can be expressed formally as a repeated strategic form game. The players in this game are the inhabitants of Santa Fe. The strategy set of the players consist of two strategies, namely go to the bar or stay home. The payoffs of the game are determined by the number of players that choose to go to the bar.

Definition 5.1 *During round t, for a bar of capacity c, the Santa Fe bar game is defined as follows:*

- $\mathcal{I} = \{1, \ldots, I\}$
- $S_i = \{\text{G}, \text{H}\}$, *where* G *means go to the bar and* H *means stay home*
- *Algorithm* $\mathring{A}$: $S \times C \to \mathbb{R}$ *measures the level of dissatisfaction of the bar attendees due to overcrowding, based on the capacity C of the bar*
- *The payoffs* $r_i(s_i, \mathring{A}(s))$ *depend on the player's strategic choice as well as the output of algorithm* $\mathring{A}$

This Santa Fe bar game is an abstraction of the so-called *uniform network game* in which user applications have two strategic choices, namely transmit or skip.

The Nash equilibrium solution of this game is a mixed strategy profile. If there are 100 inhabitants of Santa Fe ($I = 100$), and if the capacity of the bar is 60 ($c = 60$), then the Nash equilibrium strategy for all players is to go to the bar 60% of the time and to stay home 40% of the time. Like the game of Matching Pennies, the Santa Fe bar game is a game of complete information with a mixed strategy Nash equilibrium. If all players deduce this Nash equilibrium *a priori*, and if all players solemnly play this Nash equilibrium strategy at all times, then indeed play converges to the Nash equilibrium, and attendance at the bar converges to capacity. This unlikely outcome occurs only under the assumption of common knowledge of rationality.

In an attempt to find a solution to this game in the absence of common knowledge, Arthur noted the following. On a given night, the inhabitants of Santa Fe must decide whether or not to go to the bar. At one extreme, if everyone predicts that everyone else will go to the bar, then no one goes to the bar; but then, no one is happy with their decision. On the contrary, if everyone predicts that no one is going to go to the bar, then everyone will indeed go to the bar; once again, no one is happy with their decision. In either case, no one is happy; thus, neither of the above situations will persist. In particular, the conclusion is that there is no uniform *rational* behavior which maximizes overall happiness. (This outcome is reminiscent of the ecological game of Hawks and Doves in which it is also preferable to behave in contrast with the behavior of others.)

Based on these observations, Arthur simulated the behavior of *boundedly rational* agents who use inductive reasoning to build expectational models. These models are formed from a pool of simple functions which aim to predict attendance at the bar. Initially, the agents randomly select a fixed number of predictor functions from the pool: $p_1, \ldots, p_K$. All agents maintain a vector of non-negative weights associated with their predictor functions. In particular, $w_i^t(p_k)$ is the weight associated with the kth predictor function by agent i at time t. Initially, all weights are uniform. Throughout the simulation, the agents monitor the accuracy of their predictor functions and update the corresponding weights accordingly. The play of the game is prescribed by the most accurate predictor functions, which is indicated by their relative weights. Agent i utilizes the following algorithm, during round t of the game:

1. PLAY: $s_i^t = p^*(\{s^{t-1}\})$, where $w_i^{t-1}(p^*) \geq w_i^{t-1}(p_k)$, for all p_k
2. LEARN: for all predictors p_k, $w_i^t(p_k) = f(w_i^{t-1}(p_k))$, where f is *s.t.*:
   - $w_i^t(p_k) > w_i^{t-1}(p_k)$, if $p_k(\{s^{t-1}\})$ is accurate,
   - $w_i^t(p_k) < w_i^{t-1}(p_k)$, otherwise.

In this model of belief-based reasoning, agents exhibit rational play and boundedly rational learning. Their play is rational in the sense that it is optimal with respect to their beliefs, where beliefs are described by predictor functions. Their capacity for learning is bounded by their pool of simple predictor functions. Arthur simulated this repeated game and obtained an efficient solution in which the overall attendance at the bar stabilized near the capacity of the bar. The outcome of this simulation is consistent with the theoretical results reported in the last section. Moreover, the general specification of the above learning algorithm gives rise to an entire class of algorithms, some of which may satisfy merging and/or calibration, thereby converging to Nash and/or correlated equilibrium.

The Santa Fe bar game can easily be extended to a New York City bar game in which there are N bars of capacity c_n, for $1 \leq n \leq N$. This extension lends itself as an abstraction of the problem of routing network packets. In particular, links have fixed bandwidth capacities, and routers make on-line decisions as to the best possible route based on their congestion forecasts. The study of on-line learning algorithms that play such games well is the focus of this thesis research.

5.2 Congestion Game

The congestion game is a model of a more general network setting in which players choose a rate of transmission, and receive a payoff based on this rate together with the overall network delay. The game is described as follows. The possible range of transmission rates form the strategy sets $S_i = [0,1]$, for all players $i \in \mathcal{I}$. Given a vector C of link bandwidths, a prespecified routing algorithm Å: $S \times C \to \mathbb{R}$, routes the information imposing a delay of δ units. The payoffs are given by $r_i(s_i, \text{Å}(s))$; in other words, the payoffs to player i depend on i's selected transmission rate as well as the overall delay that results from congestion throughout the network. The congestion game can be further generalized by integrating it with the New York City bar game described above. In particular, players choose both the transmission rate and the route by which to transmit; the strategy sets $S_i = [0,1] \times \{1, \ldots, N\}$.

The congestion game is difficult to analyze because it does not incorporate most of the usual game-theoretic assumptions. In particular, there is no common knowledge of rationality; players do not know the others, and so cannot be sure that the others are rational. In fact, it is straightforward to interpret this game as a game against nature. Moreover, the payoff structure of the game is not commonly known; players do not necessarily know even their own payoffs. Since the congestion game provides its players with so little *a priori* information, belief-based models of this game are cumbersome and difficult to analyze. This thesis will investigate the development of learning algorithms for this and related games which weight strategies solely on the basis of performance, without requiring that players maintain prior beliefs about expected network congestion.

5.3 Free-Rider Problem

The provision of a public good, such as the internet, is a standard example of the free-rider problem in economics. Specifically, although all participants benefit from the supply of a public good, everyone prefers not to contribute to the cost of supplying the good. The free-rider problem can be analyzed as a Bayesian game of incomplete information, under the assumption that the preferences of the players are not common knowledge. On the contrary, the costs incurred by the players are private information.

1 \ 2	C	D
C	$1-c_1$, $1-c_2$	$1-c_1$, 1
D	1, $1-c_2$	0,0

Figure 9: The Free-Rider Problem

Consider the following description of the free-rider problem as a two player Bayesian game. Two players make a simultaneous decision about whether or not they wish to incur some of the cost involved in supplying a public good. The costs to the players are private knowledge, denoted c_1 and c_2, where $0 \leq c_1, c_2 \leq 1$. If neither player chooses to contribute, then the good is not provided and the payoff to both players is 0. If both players contribute, then the good is provided at a cost to them both, yielding payoffs of $1 - c_1$ and $1 - c_2$ to players 1 and 2, respectively. There are two symmetric Bayesian Nash equilibria in which the cost of the good is incurred by only one of the players, but both players benefit from its provision. The payoff matrix in Figure 9 depicts the various outcomes.

A learning algorithm such as Bayesian updating that satisfies the merging property could potentially provide a long-run solution to this game. However, this game poses an additional complication, namely the existence of multiple equilibria. It is unclear how to design a learning algorithm which gives rise to an equilibrium that alternating between these two equilibria, thereby fairly distributing the cost of supplying this public good among its beneficiaries.

6 Conclusion

The idea of learning to play equilibrium strategies in repeated games is an active area of research in the game-theoretic community. Game theorists are primarily concerned with the outcome of learning algorithms in the limit: *i.e.*, over an infinite amount of time. One of the goals of this research is to apply computer science ideology to learning theory. In particular, this thesis will consider imposing restrictions on traditional learning algorithms such that players learn to play approximations to equilibrium strategies in bounded amounts of time. The idea of such bounded learning algorithms is to quickly learn to exploit the obvious, while ignoring any subtleties.

Bounded learning is applicable to network games, in which players learn to utilize networks during times of minimal congestion. These games are atypical as compared with traditional games described in the game-theoretic literature, since their underlying structure is not commonly understood by the players, and moreover, common knowledge of rationality is not a valid assumption. As such, this class of repeated games does not naturally lend itself to belief-based learning algorithms. Rather, this thesis will investigate learning algorithms for network games that are analyzed solely on the basis of performance, without requiring that players maintain prior beliefs about expected network congestion. In sum, the initial focus of this thesis is to explore an application of computer science ideology to learning algorithms in game theory; secondly, bounded game-theoretic learning will be applied to routing and congestion problems in network environments.

References

[1] Brian Arthur. Inductive reasoning and bounded rationality. *Complexity in Economic Theory*, 84(2):406–411, May 1994.

[2] Robert Aumann. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55(1):1–18, January 1987.

[3] Robert Aumann. Game theory. In John Eatwell, Murray Milgate, and Peter Newman, editors, *Game Theory*. W.W. Norton & Company, Inc., New York, 1987.

[4] Robert Aumann and Adam Brandenburger. Epistemic conditions for Nash equilibrium. *Econometrica*, 63:1161–1180, 1995.

[5] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.

[6] Thomas Cover and Joy Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., New York, 1991.

[7] A.P. Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77(379):605–613, 1982.

[8] Ronald Fagin, Joseph Halpern, Yoram Moses, and Moshe Vardi. *Reasoning About Knowledge*. MIT Press, Cambridge, 1995.

[9] Dean Foster and Rakesh Vohra. Calibrated learning and correlated equilibrium. *Preprint*, 1995.

[10] Dean Foster and Peyton Young. Learning with hazy beliefs. *Preprint*, 1996.

[11] Yoav Freund and Robert Schapire. Game theory, on-line prediction, and boosting. In *Proceedings of the 9th Annual Conference on Computational Learning Theory*. ACM Press, May 1996.

[12] Ehud Kalai and Ehud Lehrer. Rational learning leads to Nash equilibrium. *Econometrica*, 61:1019–1045, 1993.

[13] Nick Littlestone and Manfred Warmuth. The weighted majority algorithm. *Information and Computation*, 108:212 – 261, 1994.

[14] Anatol Rapoport. *Two-Person Game Theory*. University of Michigan Press, Ann Arbor, 1966.

[15] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, New York, 1982.

[16] Ronald J. Vetter and David H.C. Du. Issues and challenges in ATM networks. *Communications of the ACM*, 38, February 1995.

[17] John von Neumann and Oskar Morgenstern. *The Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.