

# Modeling Billiards Games

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

## ABSTRACT

Two-player games of billiards, of the sort seen in recent Computer Olympiads held by the International Computer Games Association, are an emerging area with unique challenges for A.I. research. Drawing on the insight gained from our victory in the 2008 ICGA billiards tournament, we define a game-theoretic model of these types of billiards games. The modeling is surprisingly subtle. While sharing features with existing models (including stochastic games, games on a square, recursive games, and extensive form games), our model is distinct, and consequently requires novel analysis. We focus on the basic question of whether the game has an equilibrium. For finite versions of the game it is not hard to show the existence of a pure strategy Markov perfect Nash equilibrium. In the infinite case, it can be shown that under certain conditions a stationary pure strategy Markov perfect Nash equilibrium is guaranteed to exist.

## Categories and Subject Descriptors

I.2.1 [Artificial Intelligence]: Applications and Expert Systems—*games*

## General Terms

Theory

## Keywords

Billiards games, game theory, stochastic games, equilibria, best response

## 1. INTRODUCTION

The International Computer Games Association (ICGA) has in recent years introduced computer billiards as a new game in the Computer Olympiads. Our involvement began when we entered and won the 2008 ICGA computer billiards tournament. In the tournament, software agents compete against one another using a physics simulator [8]. As the game progresses in turns, an agent is presented with the state of all balls on the table, and has to then decide on an action specified by a vector of five real numbers. The five numbers describe the orientation of the cue stick in 3D space,

**Cite as:** Modeling Billiards Games, Christopher Archibald and Yoav Shoham, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. XXX–XXX.

Copyright © 2009, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

the spot on the cue ball where the cue stick will impact it, and the velocity that the cue stick will impart to the cue ball upon impact. Then, multivariate Gaussian noise is added to this chosen action (with known standard deviations), and the resulting action is input into the deterministic simulator. Some state results, and the process repeats until one of the agents has won. The precise tournament and scoring format have changed from year to year. Typically, each agent plays a series of games against each other agent, and points are earned for each victory. While still young relative to other tournaments, computer billiards is starting to attract the attention of the research community; see, e.g., Smith [16] for a description of a Computer Olympiad winning agent.

Games of billiards<sup>1</sup> have several characteristics that make them unique among games played by computer agents, and indeed among games in general. In particular, they have continuous state and action spaces, actions are taken at discrete-time intervals, there is a turn-taking structure, and the results of actions are stochastic. Computer chess, of course, features discrete state and action spaces, and actions have deterministic outcomes. Even robotic soccer, also a flourishing field, which does feature continuous state and action spaces, is different since it features continuous control and concurrent actions by both teams. Thus the challenge of billiards is novel.

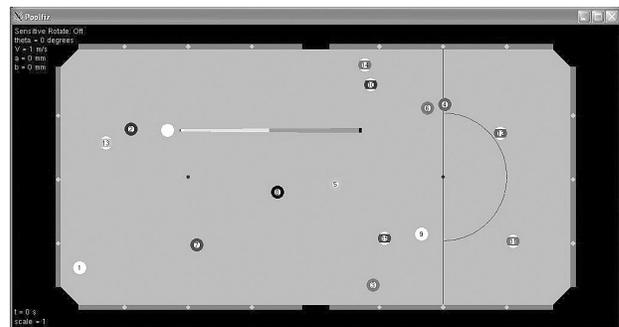


Figure 1: Display window for the ICGA billiards simulator.

Tournaments can advance science in a number of ways, and one is to put stress on existing theoretical models. The most common stress is to expose the idealizing assumptions embodied in the model. For example, the Trading

<sup>1</sup>We use the term here to refer all cue games, played with balls on a table.

Agent Competition (TAC) [18] demonstrated that game theoretic analysis does not necessarily provide insight in complex games. But at least the TAC setting is easily modeled game theoretically. Billiards games present a new kind of stress: it is not clear what formal model to apply in the first place. Game theory is blessed with a large number of different game models, some more familiar than others, and one would think that at least one of them would be appropriate. So it is something of a surprise that none are; each one is either not rich enough to apply, or too general to serve as a basis for analysis. This paper is devoted to providing a suitable model for billiard games, and proving the existence of equilibria in that model.

The paper is organized as follows. In the next section we go through various models that are suggested by features of the game and show that none of them are appropriate without change. We then provide a model, in two versions. The easy case is for finite billiard games; in this case the natural model is the extensive form. While it must be modified to allow for an uncountably infinite action space, the existence of a pure strategy Markov perfect Nash equilibrium still follows from a backward induction argument. The model for the infinite case is involved, and is based on a class of stochastic games which we term turn-taking stochastic games; these lie in between single-controller stochastic games and (single agent) MDPs. Here, with some assumptions, we can prove the existence of stationary pure strategy Markov perfect Nash equilibria. We conclude with a result on best response strategies that demonstrates the application of this model to billiard agent design.

## 2. IN SEARCH OF A MODEL

With the large number of existing game models, one would be expected to capture billiards games. Surprisingly, the unique characteristics of billiards games makes the modeling of it very subtle. Each distinct feature of billiards games suggests a match with existing game forms, some more familiar than others. In the following sections the different features of billiards games are examined along with existing game models that capture these features.

### 2.1 Continuous Action Space

One feature of billiards games is the continuous action space. The most basic of existing game models that captures this are *games on the square*, which are the continuous action space version of normal form (matrix), single stage games [13]. In these games each player chooses an action from a continuous interval, (say  $[0, 1]$ ). The payoff of the game is defined by a function  $K(a_1, a_2)$ , which specifies a real value reward for each pair of player actions.

Games on the square have been studied since the early days of game theory [2]. Early work showed that games on the square will have a value and a Nash equilibrium will exist if the function  $K$  is continuous in the joint action space. If  $K$  is discontinuous the game may still have a solution, depending upon the exact nature of the discontinuity. *Games of timing* [15] fit in this category, as there is a discontinuity in payoff based on which player acts first. Games on the square help build an understanding of continuous action spaces, but are too limited to capture other characteristics of billiards games.

### 2.2 Turn-taking

Another distinctive feature of billiards games is that only one player acts at a time. *Extensive form games* [5] are the most common class of games built upon this idea and are in this sense a great match to billiards games. Extensive form games typically consist of finite game trees, where one player makes a decision at each node, and payoffs are received at the leaf nodes. These games can be converted to normal form games, and thus have all of the typical normal form solutions, including Nash equilibria. The turn-taking nature of extensive form games also introduces new solution concepts and methods for determining solutions, both of which are instructive in the context of billiards games. A solution concept basic to extensive form games is that of a *subgame perfect equilibrium*, which is a Nash equilibrium where the strategies also form a Nash equilibrium for the game starting at any node in the game tree. This concept is applicable to billiards games. One characteristic of billiards games that is not captured by typical extensive form games is the possible infinite nature of the game. Payoffs in billiards occur only when the game stops, but there is no limit on how long this will be. This characteristic is found in the games described in the next section.

### 2.3 Reward upon Game Termination

*Recursive games*<sup>2</sup> were introduced in [3]. A recursive game is a finite set of stage games, or game elements, where the result of each stage game is a reward to the players or a probabilistic transition to a new stage game, but not both. This means that once the players receive a reward, the game ends. Recursive games as introduced by Everett do not specify the type of the stage games, and so these could be normal form games, games on the square, or any other two-player game. Everett gave conditions on the stage games that ensure existence of  $\epsilon$ -Nash equilibria. Everett's methods rely upon the set of stage games being finite and do not immediately extend to the case where the set of stage games is uncountably infinite. More recent work on recursive games [17] has maintained the restriction that the set of stage games be finite. Recursive games capture the concept of a game only having reward upon termination, but they cannot be easily adapted to represent continuous state spaces. The fact that recursive games can be viewed as a special case of stochastic games suggests a natural next place to look.

### 2.4 Continuous State Space

*Stochastic games* were introduced by Shapley [14] and have been studied extensively since. Most generally, a stochastic game consists of a set of players, a state space and an action space for each state of the game [10]. For each combination of player's actions, a payoff and subsequent probabilistic transition to another state in the state space is defined. Most commonly, stochastic games have finite state and action spaces, and so a stochastic game is simply a collection of single stage normal form games. At each discrete time step, both players take actions, receive a reward, and play transitions to a new stage game (possibly the same one). The players attempt to maximize their reward over the infinite future according to some criterion. The most common

<sup>2</sup>The authors thank Abraham Neyman for bringing this model to their attention.

criteria used are the discounted total reward, where future payoffs are discounted by some  $0 \leq \beta < 1$ , and average total reward, where players seek to maximize

$$\lim_{k \rightarrow \infty} \frac{\sum_{n=1}^k \text{Reward in stage } n}{k}.$$

Stochastic games have been studied with continuous state and action spaces [12, 9], and in this most general form, a billiards game can be represented. Much of the related literature considers only the case of discounted rewards, although some work also addresses undiscounted (positive) rewards. In each case, continuity conditions on the reward function were required to prove the existence of equilibrium strategies for the two players. The simultaneous actions of the players are the main component of the stochastic games that necessitate such conditions. In the case of billiards games, there are no simultaneous actions by the players, and so these restrictions on the reward function are unnecessary.

We get closer with *single-controller stochastic games* [10], a class of stochastic games where the actions of both players affect the payoff they receive, but only one player’s action affects the transition to the next state. However, both players are still acting simultaneously in single-controller stochastic games, and so an even more restricted setting is sought. If we limit ourselves to single player stochastic games, then we have *Markov Decision Processes (MDPs)* with their extensive literature [4]. In MDPs both the transition of the game and the payoff of the game clearly depend upon only the actions of one player. Like MDPs, billiards games consist of a single agent taking actions at discrete time intervals. The main difference is that the player taking actions will change based upon the current state. Billiards games belong in a space between single controller stochastic games and MDPs. We call games in this gap *turn-taking stochastic games*, and present billiards games as an example of a turn-taking stochastic game.

Even though stochastic games with continuous state and action spaces are not a perfect fit for billiards, they are nevertheless the only existing model that can completely represent billiards games. As previously noted, the model for these stochastic games is more general than is necessary for billiards games, as it has to account for simultaneous actions by the players and accumulation of reward at each state visited. While the results in the stochastic game literature are certainly applicable to billiards games, with a more specific model it is possible to start with fewer assumptions and produce results that are more precise. For example, in the most applicable stochastic game work, Novak [11] shows that player 2 has an optimal mixed strategy while player 1 has an  $\epsilon$ -optimal mixed strategy. Using the billiards game model we present in this paper, the existence of an equilibrium in pure strategies for both players is shown, without the same assumptions on the reward function.

While the results presented here are perhaps not surprising to those familiar with the relevant stochastic game literature, the distinctions are noteworthy, specifically when applying this model to guide understanding of billiards games and actual billiard agent design. The more specific model also allows for different, and arguably simpler, proof techniques. One purpose of this paper is to put forward a theoretical model of billiards games to enable future theoretical work on billiards games by the computer billiards community. As such, it is reasonable that future results will also be

more precise and applicable if the model used to generate them is as specific to billiards games as possible.

## 2.5 Techniques

Methods for determining solutions to these different game models inspired this paper’s approach for billiards games. Extensive form games are commonly solved using backward induction. Payoffs are specified at the leaf nodes in the game tree. Working backwards from the leaves, each decision node can be considered, and is assigned a value equal to the best choice among its children for the player who makes the decision at that node. In the case of stochastic transitions, the action with the highest expected value is selected. This process is repeated until every node in the tree has a value, and an optimal action is determined for each player’s decision node. The strategy pair where each player chooses the action that maximizes the (expected) value of the resulting node will, by construction, form a subgame perfect Nash equilibrium. This technique will inspire our approach to finite length billiards games.

Infinite length MDP models are often approached using value iteration equations and the existence of a single fixed point solution to these equations is proved. This single fixed point specifies the value of each state of the game. The optimal strategy for the agent then consists of taking the action that maximizes the expected value of the next state. Conditions under which optimal strategies for the agent will exist are presented in the MDP literature for different situations [6, 1]. We employ similar techniques in our analysis of infinite length billiards games.

## 3. A BILLIARDS GAME MODEL

We present a model for billiards games which captures all of the elements essential to understanding billiards games. As previously noted, billiards games belong to a class of turn-taking stochastic games. The model presented here is general enough to represent other games in this class, although our analysis is driven by the motivating application of billiards games.

*Definition 1.* A *Two-player Zero-sum Billiards Game*<sup>3</sup> is a tuple  $(S, A, \lambda, p, s^0, C, r)$  where:

- $S \subset \mathbb{R}^n$  is a compact  $n$ -dimensional state space, represented by a vector  $s = (s_1, \dots, s_n)$  of real numbers.
- $A \subset \mathbb{R}^m$  is the compact  $m$ -dimensional action space for players 1 and 2, where an action is represented by a vector  $a = (a_1, \dots, a_m)$  of real numbers.  $a^t \in A$  is the action chosen at time step  $t$ .
- $\lambda : S \mapsto \{1, 2\}$  is a function denoting the player whose turn it is in a given state of the game.  $\lambda(s^t)$  indicates the player whose turn it is to play in state  $s^t$ .
- $p : S \times A \mapsto \Delta(S)$  is the transition function, where  $\Delta(S)$  is the set of all probability distributions over  $S$ .
- $s^0$  is the starting state.
- $\lambda(s^0)$  is the player who gets the first turn of the game.

<sup>3</sup>This game could also be called a *Two-player zero-sum turn-taking stochastic game*, but the term *billiards game* is used here to remind of the motivation.

- $C \subseteq S$  is the set of terminating states, which is a closed subset of the state space.
- $r : C \mapsto \mathbb{R}$  is the reward function, where  $r(s)$  specifies the amount that player 2 must pay to player 1 if the game ends in state  $s \in C$ .

The game begins in state  $s^0$  and player  $\lambda(s^0)$  specifies an action. The next state,  $s^1$ , is determined from  $p(\cdot|s^0, a^0)$ , and play continues with player  $\lambda(s^1)$  specifying the next action. This cycle continues until the state is some  $s^T \in C$ , at which point the game ends, and player 1 receives  $r(s^T)$  from player 2.

### 3.1 Strategies in Billiards Games

We define, for all  $n \geq 1$ , the set of histories,  $H^n$ , as the cross product  $S \times (A \times S) \times \dots \times (A \times S)$   $[S \times n(S \times A)]^n$ . We call an element  $h^n = (s^0, a^0, s^1, \dots, a^{n-1}, s^n)$  a history. A *pure* strategy in a billiards game is a mapping  $\sigma^n : H^n \mapsto A$ . A strategy is called a *Markov strategy*, if  $\sigma^n(h^n) = \sigma^n(h'^n)$ , whenever  $h^n(s^n) = h'^n(s^n)$ . This means that at stage  $n$  of play, the player will take the same action for each state, independent of how the state was reached. A strategy  $\sigma$  is *stationary*, if it is a Markov strategy and selects the same action in each state regardless of the game's current stage. A pair of strategies  $(\sigma_1, \sigma_2)$  will induce a distribution over histories  $P_{\sigma_1, \sigma_2}^\infty$ , with which an expectation operator  $E_{\sigma_1, \sigma_2}^\infty$  can be associated, indicating the expected amount that player 2 will pay player 1 if both players follow strategies  $\sigma_1$  and  $\sigma_2$ . The strategies  $\sigma_1$  and  $\sigma_2$  form a *Nash equilibrium*, if  $E_{\sigma_1, \sigma_2}^\infty \geq E_{\sigma'_1, \sigma_2}^\infty$  for all other strategies  $\sigma'_1$  of player 1, and  $E_{\sigma_1, \sigma_2}^\infty \leq E_{\sigma_1, \sigma'_2}^\infty$  for all other strategies  $\sigma'_2$  of player 2. A Nash equilibrium is a *Markov perfect equilibrium* if both players have Markov strategies in the equilibrium and these strategies form a Nash equilibrium for any starting state of the game.

## 4. EXISTENCE OF EQUILIBRIUM

A natural first question to ask about a new model, such as the billiards game model, is what conditions ensure the existence of a Nash equilibrium in the game. The remainder of this section is focused on answering this question.

ASSUMPTION 1.  $\int f(\cdot)dp(\cdot|s, a)$  is continuous in  $A$  for any  $f \in B(S)$ , where  $B(S)$  is the set of all bounded real-valued functions on  $S$ .

Continuous functions on compact sets are guaranteed to have a maximum value and a minimum value, and this fact is used to select an action in each state, given a value function over the other states. The next lemma will assist in showing the results of this section.

LEMMA 1. If a billiards game has a value  $v^*(s), \forall s \in S$ , where  $v^*(s)$  is the unique fixed point to the value iteration equation

$$v'(s) = \begin{cases} \max_a \left[ \int_S v(\cdot)dp(\cdot|s, a) \right] & \text{if } \lambda(s) = 1 \\ \min_a \left[ \int_S v(\cdot)dp(\cdot|s, a) \right] & \text{if } \lambda(s) = 2 \end{cases}$$

then the strategies

$$\sigma_1(s) = \arg \max_a \left[ \int_S v^*(\cdot)dp(\cdot|s, a) \right]$$

and

$$\sigma_2(s) = \arg \min_a \left[ \int_S v^*(\cdot)dp(\cdot|s, a) \right]$$

will form a stationary pure strategy Markov perfect Nash equilibrium in the game.

PROOF. Clearly the two strategies are stationary and pure, as they have no dependence on the stage of the game and the arg max and arg min operators will return a single action for each state. These min and max actions will exist due to Assumption 1. If both players follow strategies  $\sigma_1$  and  $\sigma_2$ , then  $v^*(s)$  is equal to the expected total amount player 1 receives from player 2 at the end of the game starting in state  $s$ , since it is the unique fixed point of the above equation. Neither player will have any incentive to deviate, as their actions are already optimal with respect to  $v^*$ . Thus, the strategies  $\sigma_1$  and  $\sigma_2$  will be the best responses to each other, and they will form a Nash equilibrium. This equilibrium is Markov perfect since the strategies constitute a Nash equilibrium for the subgame starting in any state  $s$ .  $\square$

Conditions are now considered under which billiards games will have a value. Different modifications of billiards games are examined, where motivation for them could exist. Both the bounded and unbounded length cases are considered.

### 4.1 Bounded Length Case

We define, for each billiards game, a game length,  $K$ , which will be game specific. This will denote the maximal length of the billiards game. If play has not terminated by round  $K$ , then the game will be a tie, and player 1 will not receive any payoff from player 2. This is inspired by billiard's stalemate rule, which allows a game to be terminated by referee if progress towards a conclusion is not being made. With a limited number of balls on the table, as in typical billiards games, it is reasonable to conclude that a large enough  $K$  will not affect player's strategies significantly, assuming those strategies are designed to move the game towards a conclusion.

The maximal game length,  $K$ , allows us prove the existence of a Nash equilibrium in the game, and reason about the players strategies in that equilibrium.

THEOREM 1. If Assumption 1 holds, then a bounded length billiards game will have a pure strategy Markov perfect Nash equilibrium.

PROOF. Consider the situation in round  $K - 1$ . Due to Assumption 1, in all states  $s$  where  $\lambda(s) = 1$ , player 1 will have a clearly defined best action, which will be entirely independent of player 2's strategy, since after this move the game will end. Thus, player 1's best action choice will be  $\arg \max_a \int_S r(\cdot)dp(\cdot|s, a)$  for all states  $s$  such that  $\lambda(s) = 1$ . A similar strategy will be best for player 2 in all states where  $\lambda(s) = 2$ , replacing max with min. To each state  $s$ , then, we can assign a value in round  $K - 1$ ,

$$v^{K-1}(s) = \begin{cases} \max_a \int_C r(\cdot)dp(\cdot|s, a) & \text{if } \lambda(s) = 1 \\ \min_a \int_C r(\cdot)dp(\cdot|s, a) & \text{if } \lambda(s) = 2 \end{cases}$$

which will be the expected amount that player 2 will pay player 1 if the game is in state  $s$  during round  $K - 1$ . We can repeat this general process for each round  $K - 2, K -$

$3, \dots, 2, 1, 0$ , by iterating the following equation

$$v^{n-1}(s) = \begin{cases} \max_a [\int_C r(\cdot) dp(\cdot|s, a) \\ + \int_{S-C} v^n(\cdot) dp(\cdot|s, a)] & \text{if } \lambda(s) = 1 \\ \min_a [\int_C r(\cdot) dp(\cdot|s, a) \\ + \int_{S-C} v^n(\cdot) dp(\cdot|s, a)] & \text{if } \lambda(s) = 2 \end{cases}$$

Since in each round of the game there is a value for each state, we can apply Lemma 1 and conclude that a pure strategy Markov perfect Nash equilibrium will exist in the game. We note here that the optimal strategies from Lemma 1 are no longer stationary, because now each value corresponds to a (state,stage) pair, and thus the corresponding strategy as defined in Lemma 1 will depend on the stage of the game, making it Markov. Thus, a separate optimal strategy  $\sigma_1^{n*}(s) = \arg \max_a \int_S v^n(\cdot) dp(\cdot|s, a)$  will exist for player 1 in each stage of the game  $0 \leq n < K$ . Player 2 will have similar optimal strategies, replacing max with min.  $\square$

## 4.2 Unbounded Length Case

In this section billiards games with no bound on length are analyzed. Within this setting there are different ways of proving the existence of equilibria. Two separate situations are discussed in this section: games with discounted rewards, following the lead of the stochastic game literature, and games without discounting. The conditions under which an equilibrium will exist differ for each case, and they are considered separately in the next two sections.

### 4.2.1 Discounted Reward

In general, a discount factor denotes a preference of the players for short games over long games. There may be legitimate reasons for introducing this preference, which will depend on specific game settings. An agent's time could be valuable, in which case it would be desirable to finish the game as quickly as possible so the agent can do something else, like play more games against more opponents. The discount factor could also represent uncertainty about the future, beyond the uncertainty represented by the transition function. For example, it could represent the probability, however slight, of the game ending without cause after any round. Regardless of reason, we present here analysis of billiards games with discounted rewards.

In this case, players will choose actions to maximize their total expected discounted reward. A discount factor  $0 \leq \beta < 1$  is given. The value iteration equation will then be the following

$$v'(s) = \begin{cases} \max_a [\int_C r(\cdot) dp(\cdot|s, a) \\ + \beta \int_{S-C} v(\cdot) dp(\cdot|s, a)] & \text{if } \lambda(s) = 1 \\ \min_a [\int_C r(\cdot) dp(\cdot|s, a) \\ + \beta \int_{S-C} v(\cdot) dp(\cdot|s, a)] & \text{if } \lambda(s) = 2 \end{cases}$$

LEMMA 2. *A discounted billiards game will have a value.*

PROOF. The value iteration equation is a contraction, and as such, the Banach fixed point theorem [7] will apply, so a unique fixed point solution exists, which we denote  $v^*(\cdot)$ . Due to length considerations we do not include the full proof here. The proof follows a proof of the same fact for the infinite undiscounted case which is included in the next section. A  $\beta$  must be added into each equation and the final inequality will hold because  $\beta$  is strictly less than 1.  $\square$

Having shown that the discounted billiards game has a value, we conclude that a stationary pure strategy Markov perfect Nash equilibrium will exist.

THEOREM 2. *A stationary pure strategy Markov perfect Nash equilibrium will exist in the discounted Billiards game.*

PROOF. Direct result of lemmas 1 and 2  $\square$

### 4.2.2 Undiscounted Reward

In the most general case of billiards games, we would like to consider the undiscounted future, and determine whether a Nash equilibrium will exist in this case. To show existence of an equilibrium, we need to make a few more assumptions about the transition function and state space of billiards games, which we introduce in the following section.

### 4.2.3 The Value of Billiards

We note that the state space of a billiards game can be partitioned into a finite number of distinct subsets, where each subset is compact and corresponds to all states of the billiards game with a specific subset of the balls on the table. If a game has  $n$  balls, then there will be  $2^n$  partitions, one for each distinct subset of balls. Let  $B$  be a subset of balls, and  $S_B$  be the partition of the state space corresponding to all non-terminal arrangements of the balls in  $B$  on the table. A partial ordering over these partitions can be created, so that  $S_B \preceq S_{B'}$  if  $|B| \leq |B'|$ . We will assume that the transition function gives a probability of zero to moving from a state  $s \in S_B$  to a partition  $S_{B'}$  if  $S_B \preceq S_{B'}$ . This simply restricts our attention to billiards games where once a ball is no longer in play, it cannot later return to play. The minimal partition in this partial ordering is denoted by  $S_C = C$ , which is the set of terminal states. We will refer to the various levels of the partial ordering  $\preceq$  as  $C, C+1, C+2$ , and so on.

One additional assumption about the transition function is needed. We assume that for any state  $s \in S_B$  and any action  $a \in A$ , the probability of remaining in partition  $S_B$  is strictly less than 1. This implies that for any action, there is a non-zero probability of transitioning to a new partition of  $S$ , which, due to our previous assumption regarding the transition function, must be lower in the partial ordering. Our assumptions are summarized here.

ASSUMPTION 2. *The probability of transitioning from partition  $S_B$  to  $S_{B'}$ , where  $S_B \preceq S_{B'}$ , is 0.*

ASSUMPTION 3. *For any state  $s_b \in S_B$ , and for any action  $a \in A$ , we have that  $\int_{S_B} dp(\cdot|s_b, a) < 1$ .*

These assumptions, along with Assumption 1, allow us to show the existence of a stationary pure strategy Markov perfect Nash equilibrium in undiscounted billiards games of unbounded length. We state this result as Theorem 3.

LEMMA 3. *The value iteration equation for undiscounted billiards games is a contraction and undiscounted billiards game have a value.*

PROOF. Let  $S_{C+1}$  be a partition where only  $S_C \preceq S_{C+1}$ . By Assumption 2, for any state  $s \in S_{C+1}$  and any action  $a \in A$  there is zero probability of transitioning to any partition other than  $S_{C+1}$  and  $S_C$ . Also, the probability of ending up in  $S_C$  is positive for any  $s \in S_{C+1}$  and any action  $a \in A$ , by Assumption 3.

We consider the value iteration equation, restricted now to states  $s_{C+1} \in S_{C+1}$

$$v'(s_{C+1}) = \begin{cases} \max_a [\int_{S_C} r(\cdot) dp(\cdot|s, a) \\ + \int_{S_{C+1}} v(\cdot) dp(\cdot|s, a)] & \text{if } \lambda(s) = 1 \\ \min_a [\int_{S_C} r(\cdot) dp(\cdot|s, a) \\ + \int_{S_{C+1}} v(\cdot) dp(\cdot|s, a)] & \text{if } \lambda(s) = 2 \end{cases}$$

We show that this equation is a contraction. For any two functions  $v_1$  and  $v_2$ , both  $S_{C+1} \mapsto \mathbb{R}$ , we must show that  $\max_s |v'_1(s) - v'_2(s)| < \max_s |v_1(s) - v_2(s)|$ . We let  $\epsilon = \max_s |v_1(s) - v_2(s)|$  and consider the left-hand term. To simplify the following, attention is focused on those states  $s_{C+1} \in S_{C+1}$  where  $\lambda(s_{C+1}) = 1$ . A similar treatment works for the other states ( $\lambda(s_{C+1}) = 2$ ) as well.

$$\begin{aligned} |v'_1(s) - v'_2(s)| &= \\ & \left| \max_a \left[ \int_{S_C} r(\cdot) dp(\cdot|s, a) + \int_{S_{C+1}} v_1(\cdot) dp(\cdot|s, a) \right] \right. \\ & \quad \left. - \max_a \left[ \int_{S_C} r(\cdot) dp(\cdot|s, a) + \int_{S_{C+1}} v_2(\cdot) dp(\cdot|s, a) \right] \right| \\ & \leq \left| \max_a \left[ \int_{S_C} r(\cdot) dp(\cdot|s, a) + \int_{S_{C+1}} v_1(\cdot) dp(\cdot|s, a) \right] \right. \\ & \quad \left. - \int_{S_C} r(\cdot) dp(\cdot|s, a) - \int_{S_{C+1}} v_2(\cdot) dp(\cdot|s, a) \right| \\ & \leq \max_a \left| \int_{S_{C+1}} (v_1(\cdot) - v_2(\cdot)) dp(\cdot|s, a) \right| \\ & \leq \max_a \int_{S_{C+1}} |v_1(\cdot) - v_2(\cdot)| dp(\cdot|s, a) \\ & \leq \max_a \int_{S_{C+1}} \epsilon \cdot dp(\cdot|s, a) \\ & = \epsilon \cdot \max_a \int_{S_{C+1}} dp(\cdot|s, a) \\ & < \epsilon \end{aligned}$$

The last step is a result of Assumption 3. Thus, the value iteration equation is a contraction. Since by assumption  $S_B$  is compact, the Banach fixed point theorem [7] can be applied and we conclude that there will be a unique fixed point, which we will denote here as  $v_{C+1}^*$ .

This same technique can be used to prove the existence of a value for all other partitions from which play can transition only to themselves or to terminal partitions. Thus, for each other partition  $S'_{C+1}$  at the  $C+1$  level of the partial ordering  $\preceq$ , we have a value function  $v_{C+1}^*$ .

Consider now another partition  $S_{C+2}$  from which transitions can occur either into a terminal partition  $S_C$  or into the partition level just discussed ( $C+1$ ). The same proof technique can be repeated to show that there will be a fixed point value for all states within this partition, and all other partitions at the  $C+2$  level in the partial ordering. The main difference in the proof would be the addition of another term in the value iteration equation corresponding to the probability of transitioning to each partition at level  $C+1$ . This process can be repeated for all levels of the hierarchy,  $C+3, C+4, \dots$ , until we have a value for each state  $s \in S$ .  $\square$

**THEOREM 3.** *If the transition function obeys Assumptions 1-3, then undiscounted billiards games have a value, and a stationary pure strategy Markov perfect Nash equilibrium exists.*

PROOF. Direct result of Lemmas 1 and 3.  $\square$

## 5. BEST RESPONSE

Another natural use of a model is to give theoretical support for decisions made while designing an agent to participate in the setting described by the model. In addition to the knowledge that an equilibrium exists in the game, as shown in the previous section, knowledge of the type of strategy necessary to best respond to an opponent's strategy can motivate and justify the design decision to focus on a narrower class of strategy types. To this end, the next theorem is presented.

**THEOREM 4.** *In an undiscounted, unbounded length billiards game, if the transition function obeys Assumptions 1-3, and player 1 is playing a pure stationary strategy, then player 2's best response will also be a pure stationary strategy.*

PROOF. Let  $D : S \mapsto A$  represent player 1's stationary pure strategy, selecting action  $D(s)$  each time state  $s$  is visited. We can replace the value iteration equation with the following

$$v'(s) = \begin{cases} \int_{S_C} r(\cdot) dp(\cdot|s, D(s)) \\ + \int_S v(\cdot) dp(\cdot|s, D(s)) & \text{if } \lambda(s) = 1 \\ \min_a \left[ \int_{S_C} r(\cdot) dp(\cdot|s, a) \right. \\ \quad \left. + \int_S v(\cdot) dp(\cdot|s, a) \right] & \text{if } \lambda(s) = 2 \end{cases}$$

Since the situation is the same as for Lemma 3, we merely have to show that the  $\lambda(s) = 1$  portion of the equation is a contraction. As before we must show that for any two functions  $v_1$  and  $v_2$ , both  $S_{C+1} \mapsto \mathbb{R}$ , it is the case that  $\max_s |v'_1(s) - v'_2(s)| < \max_s |v_1(s) - v_2(s)|$ . We let  $\epsilon = \max_s |v_1(s) - v_2(s)|$  and consider the left-hand term.

$$\begin{aligned} |v'_1(s) - v'_2(s)| &= \\ & \left| \int_{S_C} r(\cdot) dp(\cdot|s, D(s)) + \int_{S_{C+1}} v_1(\cdot) dp(\cdot|s, D(s)) \right. \\ & \quad \left. - \left[ \int_{S_C} r(\cdot) dp(\cdot|s, D(s)) + \int_{S_{C+1}} v_2(\cdot) dp(\cdot|s, D(s)) \right] \right| \\ & = \left| \int_{S_{C+1}} (v_1(\cdot) - v_2(\cdot)) dp(\cdot|s, D(s)) \right| \\ & \leq \left| \int_{S_{C+1}} \epsilon \cdot dp(\cdot|s, D(s)) \right| \\ & = \epsilon \cdot \left| \int_{S_{C+1}} dp(\cdot|s, D(s)) \right| \\ & < \epsilon \end{aligned}$$

We again used the definition of  $\epsilon = \max_s |v_1(s) - v_2(s)|$ , as well as Assumption 3 in the last step. Since the value iteration equation is again a contraction, it is evident, as in Lemma 3, that each state of the billiards game will have a value, and we can conclude, as in Lemma 1, that a stationary pure strategy for player 2 will perform optimally with regard to this value function. Thus, player 2's best response to player 1's strategy will also be a pure stationary strategy.  $\square$

This result can guide as well as support design decisions made while creating agents to compete in computer pool. Suppose it is known that the opponent's agent chooses the

same action in each state, as most existing billiards agents do, thus employing a stationary strategy. Then restricting one's design space to pure stationary strategies in response does not eliminate the best response to the opponent's strategy.

## 6. CONCLUDING REMARKS

We presented a model that represents billiards games, along with initial analysis of these games under different modeling conditions and assumptions. We showed the existence of a Markov perfect Nash equilibrium in each case. In the case most closely corresponding to the actual billiards games of the ICGA computer billiards tournament, we found that the best response to a pure stationary strategy is also a pure stationary strategy. In the future less stringent requirements under which equilibria exist will be explored, as well as other solution concepts. Our model suggests many interesting experimental topics, such as determining or approximating the value in billiards games states, and computing an optimal strategy efficiently given such a value function or opponent strategy. This paper is intended to lay a foundation for a formal understanding of billiards games, (and more generally, turn-taking stochastic games), and lead to an improved comprehension of the fundamental issues in these games. This improved understanding will help advance the state of the art in billiards software agents.

## 7. ACKNOWLEDGMENTS

This work was supported by NSF grants IIS-0205633-001 and SES-0527650.

## 8. REFERENCES

- [1] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete Time Case*, volume 139 of *Mathematics in Science and Engineering*. Academic Press, 1978.
- [2] H. F. Bohnenblust, S. Karlin, and L. S. Shapley. *Games with Continuous Convex Payoff*. Annals of Mathematics Studies 24. Princeton University Press, 1950.
- [3] H. Everett. *Recursive Games*, pages 47–78. Annals of Mathematics Studies 39. Princeton University Press, 1957.
- [4] E. A. Feinberg and A. Shwartz, editors. *Handbook of Markov Decision Processes: Methods and Applications*. Kluwer Academic, 2002.
- [5] D. Fudenberg and J. Tirole. *Game Theory*. The MIT Press, 2000.
- [6] O. Hernandez-Lerma. *Adaptive Markov Control Processes*. Springer-Verlag, 1989.
- [7] W. A. Kirk and B. Sims. *Handbook of Metric Fixed Point Theory*. Kluwer Academic, 2001.
- [8] W. Leckie and M. Greenspan. *An Event-Based Pool Physics Simulator*, pages 247–262. Lecture Notes in Computer Science. Springer, 2006.
- [9] A. Maitra and T. Parthasarathy. On stochastic games 1. *Journal of Optimization Theory and Applications*, 5:289–300, 1970.
- [10] A. Neyman and S. Sorin. *Stochastic Games and Applications*. Kluwer Academic Publishers, Dordrecht, 2003.
- [11] A. S. Nowak. *Zero-sum Stochastic Games with Borel State Spaces*. Kluwer Academic Publishers, Dordrecht, 2003.
- [12] A. S. Nowak and E. Altman.  $\epsilon$ -equilibria for stochastic games with uncountable state space and unbounded costs. *SIAM J. Control Optim.*, 40(6):1821–1839, 2002.
- [13] G. Owen. *Game Theory*. Academic Press, 2001.
- [14] L. S. Shapley. Stochastic games. *Mathematics*, 39:1095–1100, 1953.
- [15] M. Shiffman. *Games of Timing*. Annals of Mathematics Studies 28. Princeton University Press, 1953.
- [16] M. Smith. Running the table: An AI for computer billiards. In *Proceedings of 21st Nat. Conf. on Artificial Intelligence*, 2006.
- [17] F. Thuijsman. *Recursive Games*, volume 570 of *NATO Science Series C, Mathematical and Physical Sciences*, chapter 16, pages 253–264. Kluwer Academic Publishers, Dordrecht, 2003.
- [18] M. P. Wellman, A. Greenwald, and P. Stone. *Autonomous Bidding Agents: Strategies and Lessons from the Trading Agent Competition*. MIT Press, 2007.