THE COMPLEXITY OF STOCHASTIC GAMES

by

Anne Condon

Computer Sciences Technical Report #863

July 1989

The Complexity of Stochastic Games

Anne Condon[1]

Computer Sciences Department
University of Wisconsin-Madison
1210 West Dayton Street
Madison, WI 53706

Abstract

We consider the complexity of stochastic games – simple games of chance played by two players. We show that the problem of deciding which player has the greatest chance of winning the game is in the class $NP \cap co\text{-}NP$.

1 Introduction

We consider the complexity of a natural combinatorial problem, that of deciding the outcome of a special kind of stochastic game. A *simple stochastic game (SSG)* is a directed graph with three types of vertices, called *max*, *min* and *average* vertices. There is a special start vertex and two special sink vertices, called the *0-sink* and the *1-sink*. For simplicity, we assume that all vertices have exactly two (not necessarily distinct) neighbors, except for the sink vertices, which have no neighbors.

The graph models a game between two players, 0 and 1. In the game, a token is initially placed on the start vertex, and at each step of the game the token is moved from a vertex to one of its neighbors, according to the following rules: At a min vertex, player 0 chooses the neighbor to which the token is moved. At a max vertex, player 1 chooses the neighbor to which the token is moved. At an average vertex, a coin is tossed to determine where the token is moved, so that it is moved to each neighbor of the average vertex with probability 1/2.

The game ends when the token reaches a sink vertex; player 1 wins if it reaches the 1-sink vertex and player 0 wins otherwise, that is, if the token reaches the 0-sink vertex or if the game never halts. The reason for the names *max* and *min* vertices is that at the max vertices, player 1 chooses its move so as to maximize the probability of eventually reaching the 1-sink vertex, whereas at the min vertices, player 0 chooses its move so as to minimize the probability of reaching the 1-sink vertex.

Informally, a *strategy* for player 0 (1) is a rule that defines what move the player takes whenever the token is at a min (max) vertex. There are two natural questions to ask about a simple stochastic game:

(i) what are the best strategies of the players; and

(ii) what is the probability that player 1 wins the game, if both players use their best strategies.

Although many algorithms that answer these questions have been studied, none run in polynomial time. Thus the natural decision problem associated with these questions - that of deciding if the probability that player 1 wins is greater than 1/2 - is not known to be in P . The purpose of this paper is not to describe new algorithms for this problem, but instead to formulate the important questions in a complexity-theoretic framework. We show that this decision problem is in $NP \cap co\text{-}NP$. Although many number theoretic problems not known to be in P lie in the class $NP \cap$ co-NP, combinatorial problems

[1]Supported by National Science Foundation grant number DCR-8802736.

that lie between P and $NP \cap$ co-NP are rare. The stochastic game problem defined in this paper is an interesting example of a simple combinatorial problem with this property.

The study of stochastic games was initiated by Shapley [10] in 1953 and many variations of the model have been investigated since then (see Peters and Vrieze [8] for a survey). The class of stochastic games studied in this paper is somewhat similar to the stochastic games with switching transitions, studied by Filar [3]. We were motivated to study these simple stochastic games while considering the power of the following complexity model: space bounded alternating Turing machines, generalized to allow random as well as universal and existential moves. (Alternating Turing machines are described in [1]). These machines model some interesting problems in optimization; a polynomial time algorithm for the simple stochastic game problem would yield polynomial time algorithms for other optimization problems as well. It would also imply that adding randomness to space bounded alternating Turing machines does not increase the class of languages they accept.

In Section 2, we give a more precise definition of the stochastic game problem considered in the paper and describe properties of stochastic games. In Section 3, we show that the decision problem associated with these questions lies in the class *NP* $\cap$ *co-NP*. In Section 4, we briefly describe variations of the *SSG* model and examine their complexity, and list some open problems.

2 Definitions and Properties of Simple Stochastic Games

In this section, we define precisely the model of stochastic games considered in this paper. We describe some basic properties of strategies of the players of such games. Many of the results presented here have previously appeared in the literature of stochastic games; we include them for completeness and because in some cases they can be simplified or strengthened for the special class of games studied in this paper.

2.1 Definitions

A *simple stochastic game (SSG)* is a directed graph $G = (V, E)$ with the following properties. The vertex set V is the union of disjoint sets $V_{max}, V_{min}, V_{average}$, called max, min and average vertices, together with two special vertices, called the *0-sink* and the *1-sink*. One vertex of V is called the *start* vertex. Each vertex of V has two outgoing edges, except the sink vertices, which have no outgoing edges. Without loss of generality, assume that the vertices of G are numbered $1, 2, \ldots n$, with $n-1$ and n being the 0- and 1-sink vertices, respectively. An example of a *SSG* is given in Figure 1

Associated with the game are two players, 0 and 1. A *strategy* τ of player 0 is a set of edges of E, each with its left end at a min vertex such that for each min vertex i there is exactly one edge (i, j) in τ. Informally, if $(i, j) \in \tau$ then in a game where player 0 uses strategy τ, the token is always moved from vertex i to vertex j. Similarly, a strategy σ of player 1 is a set of edges of E, each with its left end at a max vertex such that for each max vertex i there is exactly one edge (i, j) in σ. In the game theory literature, strategies satisfying this definition are called *pure stationary* strategies because (i) the players do not use probabilistic choice in choosing a move, and (ii) each player chooses the same move from a vertex every time that vertex is reached. We only consider pure stationary strategies in this paper because, as we will see later, both players of a *SSG* have optimal strategies of this type. (For a discussion of other types of strategies, see Peters and Vrieze [8]).

Corresponding to strategy σ is a graph G_σ, which is the subgraph of G obtained by removing from

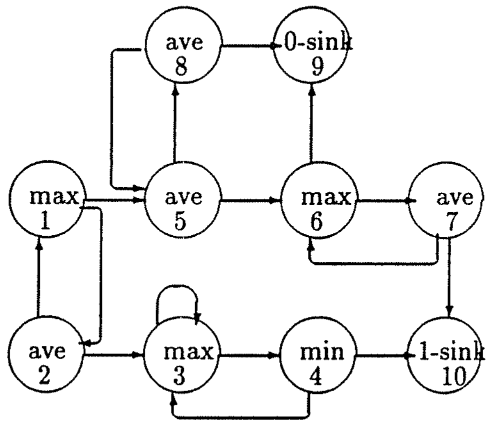


Figure 1: *A Simple Stochastic Game with 10 vertices. Vertex* 1 *is the start vertex, vertex* 9 *the* 0*-sink and vertex* 10 *the* 1*-sink.*

each max vertex the outgoing edge that is not in the strategy σ. Similarly, corresponding to a pair of strategies σ and τ, is a graph $G_{\sigma,\tau}$ obtained from G_σ by removing from each min vertex the outgoing edge that is not in τ. In $G_{\sigma,\tau}$, every max and min vertex has one outgoing edge. Thus $G_{\sigma,\tau}$ can be considered as a Markov process where the states of the process are the vertices of G and the transition probabilities $p_{ij}, 1 \leq i,j \leq n$ are defined as follows. If $i \leq n-1$ then $p_{ij} = \frac{1}{2}$ if i is an average vertex with outgoing edge (i,j); $p_{ij} = 1$ if i is a max or min vertex with outgoing edge (i,j) and $p_{ij} = 0$ otherwise. Since $n-1$ and n are sink states, we define $p_{nn} = p_{n-1n-1} = 1$, $p_{nj} = 0$ if $j \neq n$ and $p_{n-1j} = 0$ if $j \neq n-1$.

We define the *value* $v_{\sigma,\tau}(i)$ of each vertex i of G with respect to strategies σ and τ to be the probability that player 1 wins the game if the start vertex is i and the players use strategies σ and τ. That is, the value $v_{\sigma,\tau}(i)$ is the probability of reaching the 1-sink vertex from the start vertex on a random walk of the Markov process $G_{\sigma,\tau}$, starting from vertex i. In Figure 2, the subgraph of the *SSG* of Figure 1 is given, for strategies $\sigma = \{(1,5),(3,4),(6,7)\}$ and $\tau = \{(4,3)\}$ of player 1 and 0, respectively.

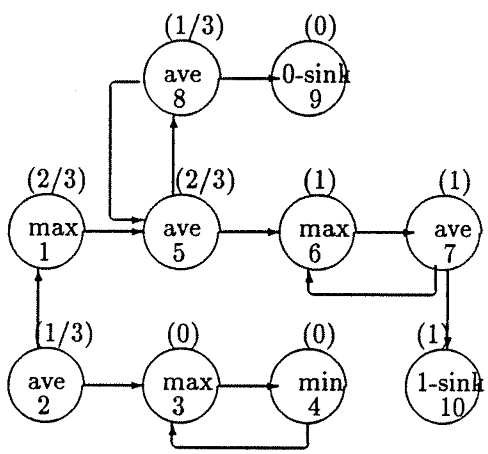


Figure 2: *Subgraph of the* SSG *of Figure 1, for strategies* $\sigma = \{(1,5),(3,4),(6,7)\}$ *and* $\tau = \{(4,3)\}$ *of player* 1 *and* 0*, respectively. The value* $v_{\sigma,\tau}(i)$ *of each vertex* i *is given above the vertex.*

Let vertices $1,\ldots,t$ be the vertices of $G_{\sigma,\tau}$ with a path to the 1-sink vertex, excluding the 1-sink vertex itself. Then clearly the values $v_{\sigma,\tau}(i), 1 \leq i \leq n$ satisfy the following conditions: If $t < i \leq n-1$ then $v_{\sigma,\tau}(i) = 0$ and $v_{\sigma,\tau}(n) = 1$; otherwise

$$v_{\sigma,\tau}(i) = \begin{cases} \frac{1}{2}(v_{\sigma,\tau}(j) + v_{\sigma,\tau}(k)), & \text{if } i \text{ is an average vertex of } G_{\sigma,\tau} \\ & \text{with outgoing edges } (i,j),\ (i,k), \\ v_{\sigma,\tau}(j), & \text{if } i \text{ is a max or a min vertex of } G_{\sigma,\tau} \\ & \text{with outgoing edge } (i,j). \end{cases} \tag{1}$$

We next show that the equations (1) have a unique solution. Let $\bar{v}_{\sigma,\tau} = (v_{\sigma,\tau}(1), \ldots, v_{\sigma,\tau}(t))$.

Lemma 1 *There is an $t \times t$ matrix Q and a t-vector $\bar{b}$ with entries in $\{0, 1/2, 1\}$ such that $\bar{v}_{\sigma,\tau}$ is the unique solution to the equation $\bar{v}_{\sigma,\tau} = Q\bar{v}_{\sigma,\tau} + \bar{b}$. Also $I - Q$ is invertible, all entries of $(I-Q)^{-1}$ are non-negative and the entries along the diagonal are strictly positive.*

Proof: Substituting 0 for $v_{\sigma,\tau}(i)$, $t+1 \leq i \leq n-1$ and 1 for $v_{\sigma,\tau}(n)$ in equations (1) and rearranging the terms in the equations, it follows that

$$\bar{v}_{\sigma,\tau} = Q\bar{v}_{\sigma,\tau} + \bar{b},$$

where $\bar{b}$ is a constant vector and Q is the the 1-step transition matrix of the vertices $\{1, \ldots, t\}$. For each i, the ith component of $\bar{b}$ is from the set $\{0, 1/2, 1\}$ and equals the probability of reaching the 1-sink vertex from i in 1 step. Also all entries of Q are in $\{0, 1/2, 1\}$. The equation $\bar{v}_{\sigma,\tau} = Q\bar{v}_{\sigma,\tau} + \bar{b}$ has a unique solution if and only if $(I - Q)$ is invertible.

To prove that $(I - Q)$ is invertible, we first argue that $\lim_{l\to\infty} Q^l = 0$. We show that for any $m \geq 1$, the sum of the terms in any row of Q^{mn} is at most $(1 - 1/2^n)^m$. Since all terms of Q^l are non-negative for any l, it follows from this that as $l \to \infty$, $Q^l \to 0$. Note that the ijth entry of Q^{mn} is the probability of reaching vertex j from vertex i in exactly mn steps. When m is 1, the sum of the terms of the ith row of Q^{mn} is at most 1 minus the probability of reaching the 1-sink vertex of $G_{\sigma,\tau}$ from i in n steps. Since the value of i is > 0, there must be a path of length $\leq n$ in $G_{\sigma,\tau}$ from i to the 1-sink. Hence the probability of reaching the 1-sink from i in n steps is at least $1/2^n$. Therefore, the sum of the terms in any row of Q^n is at most $1 - 1/2^n$. The proof for $m > 1$ is an easy induction argument.

From this, Kemeny and Snell [6] give a simple argument to show that $I - Q$ is invertible. From simple rules of algebra,

$$I - Q^l = (I - Q)(I + Q + Q^2 + \ldots + Q^{l-1});$$

taking the limit of both sides as $l \to \infty$ yields $I = (I - Q)(I + Q + \ldots + Q^l + \ldots)$. Since the left side has determinant 1, the determinant of the right side must also be 1, implying that $I - Q$ has non-zero determinant. Hence $I - Q$ is invertible. Also, $(I - Q)^{-1} = I + Q + Q^2 + \ldots$; hence every entry of $(I - Q)^{-1}$ is non-negative and the entries along the diagonal are strictly positive. □

We define the *value of the game* G to be

$$\max_{\sigma} \min_{\tau} v_{\sigma,\tau}(start).$$

Informally, the value of the game is the maximum probability that player 1 wins if it reveals its best strategy to player 0 at the start of the game; and player 0 plays its best strategy against the strategy chosen by player 1.

Given a *SSG*, a natural question is: what is its value? A related problem is to find the best strategies of the players, that is, the strategies σ and τ such that the value of the game equals $v_{\sigma,\tau}(start)$. To investigate the complexity of these problems, we consider the following decision problem for *SSG*'s.

The **SSG value problem** is: Given a SSG, is its value $> 1/2$?

In Section 3, we show that this decision problem is in the class $NP \cap co\text{-}NP$. To build up to the proof, we describe many interesting properties of simple stochastic games in Section 2.2.

2.2 Properties of Simple Stochastic Games

We next describe some properties of the value and of the strategies of simple stochastic games. We first show that the value of a *SSG* is a rational number of the form p/q where $0 \leq p, q \leq 4^n$. From this, it is possible to bound the value of a *SSG* away from 1/2, if it is not equal to 1/2. In the succeeding lemmas, we build up to the proof that both players possess "optimal" strategies – strategies that are guaranteed to ensure that the outcome of the game is the best possible for that player, regardless of what the other player does or what the start vertex is. Finally in Lemma 6 we describe a strong form of the minimax theorem for simple stochastic games.

Lemma 2 *The value of a simple stochastic game G with n vertices is of the form p/q where p and q are integers, $0 \leq p, q \leq 4^n$.*

Proof: The proof of this lemma is very similar to a proof by Gill ([4], Lemma 6.6). Let σ and τ be arbitrary strategies of players 0 and 1 respectively. Then the value of 0-valued vertices and the 1-sink vertex with respect to strategies σ and τ can be written as 0/1 and 1/1 respectively and so are clearly of the form given in the lemma. Without loss of generality, assume that all of the vertices with positive value, excluding the 1-sink vertex, are numbered from 1 to t and let $\bar{v}_{\sigma,\tau} = (v_{\sigma,\tau}(1), \ldots, v_{\sigma,\tau}(t))$. In Lemma 1 it was shown that $\bar{v}_{\sigma,\tau}$ is the unique solution to the equation $2(I - Q)\bar{v}_{\sigma,\tau} = 2\bar{b}$, where $\bar{b}$ is a constant vector and $I - Q$ is a matrix with non-zero determinant. Furthermore, the entries in $2(I - Q)$ and $2\bar{b}$ are from the set $\{0, \pm 1, \pm 2\}$.

By Cramer's rule, the value of a vertex $i \in \{1, 2, \ldots t\}$ can be represented as the quotient of two integers N_i/D, where D is the determinant of $2(I - Q)$ and N_i is the determinant of the matrix obtained by replacing the *ith* column of $2(I - Q)$ by $2\bar{b}$. The determinant of $2(I - Q)$ is at most 4^n. This follows from expansion by minors, the fact that the entries of $2(I - Q)$ are from the set $\{0, \pm 1, \pm 2\}$ and that each row has a constant number of non-zero entries, whose absolute values sum to at most 4. Hence D is at most 4^n. Also, since the value of each vertex is at most 1, the value of each N_i is an integer $\leq D$. Thus the value of each vertex of $G_{\sigma,\tau}$ is the quotient of two numbers, each of value at most 4^n. In particular this is true for the start vertex of $G_{\sigma,\tau}$.

Since the strategies σ and τ were chosen arbitrarily, $\max_{\sigma} \min_{\tau} v_{\sigma,\tau}(start)$ must be the quotient of two numbers, each of value at most 4^n. □

Lemma 3 *If the value of a simple stochastic game with n vertices is $> 1/2$, then it is $\geq 1/2 + 1/4^n$.*

Proof: Immediate from the last lemma. □

In the next lemma, it is shown that for each strategy σ of player 1, there is a strategy $\tau(\sigma)$ of player 0 that is "optimal" with respect to σ in that for every min vertex i of G, the value of i with respect to strategies $\sigma, \tau(\sigma)$ equals the minimum of the values of its children. We call $\tau(\sigma)$ the optimal strategy of player 0 with respect to strategy σ.

Lemma 4 (Howard [5]) *Let G be a simple stochastic game with n vertices and let σ be any strategy of player 1. Then there is some strategy $\tau(\sigma)$ such that for each vertex $i \in V_{min}$ with neighbors j and k,*

$$v_{\sigma,\tau(\sigma)}(i) = \min[v_{\sigma,\tau(\sigma)}(j), v_{\sigma,\tau(\sigma)}(k)].$$

Proof: Howard [5] described an algorithm to construct $\tau(\sigma)$ satisfying the lemma. The algorithm proceeds in iterations. There is a current strategy $\tau(r)$ for each iteration r; the current strategy of the initial iteration is chosen arbitrarily and is denoted by $\tau(0)$. Strategy $\tau(r+1)$ is constructed from strategy $\tau(r)$ by finding any one vertex $i \in V_{min}$ with neighbors j, k such that $(i,j) \in \tau(r)$ but $v_{\sigma,\tau(r)}(k) < v_{\sigma,\tau(r)}(j)$ and replacing (i,j) by (i,k). If no such vertex exists, the algorithm halts. The current strategy of the final iteration is $\tau(\sigma)$. From the construction of $\tau(\sigma)$, it must be the case that for each vertex $i \in V_{min}$ with neighbors j and k, $v_{\sigma,\tau(\sigma)}(i) = \min[v_{\sigma,\tau(\sigma)}(j), v_{\sigma,\tau(\sigma)}(k)]$.

To prove that Howard's algorithm is correct, it remains to show that it always halts. To simplify the notation in the proof, let $v_r(i) = v_{\sigma,\tau(r)}(i)$. The fact that Howard's algorithm halts follows from the following property, which we will prove: if the algorithm does not halt at round r then strategy $\tau(r+1)$ improves strategy $\tau(r)$ in that for each vertex l of G, $v_{r+1}(l) \le v_r(l)$ and for at least one vertex the inequality is strict. This ensures that no strategy can be repeated in the sequence $\tau(0), \tau(1), \ldots, \tau(\sigma)$ and hence the sequence is finite.

Clearly, if $v_{r+1}(l) = 0$ then $v_{r+1}(l) \le v_r(l)$ and also for n, the 1-sink vertex, $1 = v_{r+1}(n) \le v_r(n) = 1$. Therefore we restrict our attention to the vertices other than the sink vertex for which $v_{r+1}(l) > 0$. Without loss of generality let these vertices be numbered 1 through t. Let $\bar{v}_r = (v_r(1), \ldots, v_r(t))^T$. As in Lemma 1 let matrices Q_r, Q_{r+1} and vectors $\bar{b}_r$ and $\bar{b}_{r+1}$ be such that $\bar{v}_r = Q_r\bar{v}_r + \bar{b}_r$ and $\bar{v}_{r+1} = Q_{r+1}\bar{v}_{r+1} + \bar{b}_{r+1}$. Let $\bar{\Delta} = \bar{v}_r - \bar{v}_{r+1}$. We show that $\bar{\Delta} \ge 0$ and that some entry is actually > 0. Adding and subtracting $Q_{r+1}\bar{v}_r + \bar{b}_{r+1}$ to $\bar{\Delta}$ we see that

$$\bar{\Delta} = (Q_r\bar{v}_r + \bar{b}_r) - (Q_{r+1}\bar{v}_r + \bar{b}_{r+1}) + (Q_{r+1}\bar{v}_r + \bar{b}_{r+1}) - (Q_{r+1}\bar{v}_{r+1} + \bar{b}_{r+1}).$$

If $\bar{\delta} = (Q_r\bar{v}_r + \bar{b}_r) - (Q_{r+1}\bar{v}_r + \bar{b}_{r+1})$, then $\bar{\Delta} = Q_{r+1}\bar{\Delta} + \bar{\delta}$. Since Q_{r+1} is the 1-step transition matrix of vertices $1, \ldots, t$ of $G_{\sigma,\tau(r+1)}$, from Lemma 1 it follows that the matrix $(I - Q_{r+1})$ is invertible. Hence $\bar{\Delta} = (I - Q_{r+1})^{-1}\bar{\delta}$. Also from Lemma 1, all the entries of $(I - Q_{r+1})^{-1}$ are ≥ 0 and the entries along the diagonal are > 0. Therefore, it remains to show that $\bar{\delta} \ge 0$ and that one entry is actually > 0. Suppose edge (i,k) replaces (i,j) in constructing $\tau(r+1)$ from $\tau(r)$. Hence Q_r and Q_{r+1} differ in at most the ith row and similarly $\bar{b}_r$ and $\bar{b}_{r+1}$ differ in at most the ith row. Thus every entry of $\bar{\delta}$ must be 0, except possibly the ith entry.

Moreover, the ith entries of $Q_r\bar{v}_r + \bar{b}_r$ and $Q_{r+1}\bar{v}_r + \bar{b}_{r+1}$ are $v_r(j)$ and $v_r(k)$ respectively; hence the ith entry of $\bar{\delta}$ is $v_r(j) - v_r(k)$. Since edge (i,k) replaces (i,j) in constructing $\tau(r+1)$ from $\tau(r)$, it must be the case that $v_r(j) > v_r(k)$ and hence the ith entry of $\bar{\delta}$ must be > 0. This completes the proof that Howard's algorithm halts. □

The next lemma shows that in any *SSG* G, player 1 has a strategy σ' which is optimal in the following sense. Suppose player 0 uses some optimal strategy $\tau(\sigma')$ with respect to strategy σ' of player 1. Then for each vertex $i \in V_{max}$, the value of i with respect to strategies σ' and $\tau(\sigma')$ is the maximum of the values of its children. We call any strategy σ' satisfying Lemma 5 an optimal strategy of player 1. The proof of this lemma is a straightforward extension of Lemma 4; similar proofs for other classes of stochastic games are surveyed in [8].

Lemma 5 *Let G be a simple stochastic game with n vertices. Then there is a strategy σ' of player 1 such that, if $\tau' = \tau(\sigma')$ is the optimal strategy of player 0 with respect to strategy σ' of player 1, then for all vertices $i \in V_{max}$ with neighbors j and k,*

$$v_{\sigma',\tau'}(i) = \max[v_{\sigma',\tau'}(j), v_{\sigma',\tau'}(k)].$$

Proof: Again the strategy σ' can be found by an iterative method similar to Lemma 4. In the iterative algorithm to construct σ', there is a current strategy $\sigma(r)$ for each iteration r; the current strategy for the initial iteration is chosen arbitrarily. Let $\tau(r)$ denote the optimal strategy of player 0 with respect to strategy $\sigma(r)$ and let $v_r(l)$ denote $v_{\sigma(r),\tau(r)}(l)$. Strategy $\sigma(r+1)$ is constructed from strategy $\sigma(r)$ by finding any one vertex $i \in V_{max}$ such that $(i,j) \in \sigma(r)$ but $v_r(k) > v_r(j)$ and replacing (i,j) by (i,k). If no such vertex exists the algorithm halts; the current strategy of the final iteration is σ'. Let τ' be the optimal strategy of player 0 with respect to σ'. From this construction, it must be the case that for each vertex $i \in V_{max}$ with neighbors j and k, $v_{\sigma',\tau'}(i) = \max[v_{\sigma',\tau'}(j), v_{\sigma',\tau'}(k)]$.

As in Lemma 4, to show that the iterative algorithm halts, it is sufficient to show that $v_{r+1}(l) \geq v_r(l)$ and that the inequality is strict for some l. The proof of this is complicated by the fact that at each iteration, both $\sigma(r)$ and $\tau(r)$ are being changed, whereas in the construction of Lemma 4 only one strategy is changed. However, the key elements in the proof are similar.

Assume that $v_{r+1}(l) > 0$ for $1 \leq l \leq t$ and $v_{r+1}(l) = 0$ for $t < l < n$. Also, suppose that (i,k) replaces (i,j) in the construction of $\sigma(r+1)$ from $\sigma(r)$ and that (i',k') replaces (i',j') in the construction of $\tau(r+1)$ from $\tau(r)$. We restrict the proof to the case where $\tau(r+1)$ and $\tau(r)$ differ in only 1 edge; the proof can easily be generalized to the case when they differ in more than one edge.

First, consider the vertices $1,\ldots,t$. Let $\bar{v}_r = (v_r(1),\ldots v_r(t))^T$, $\bar{v}_{r+1} = (v_{r+1}(1),\ldots v_{r+1}(t))^T$, $\bar{v}_r = Q_r\bar{v}_r + \bar{b}_r$ and $\bar{v}_{r+1} = Q_{r+1}\bar{v}_{r+1} + \bar{b}_{r+1}$ for some $Q_r, Q_{r+1}, \bar{b}_r$ and $\bar{b}_{r+1}$. As in Lemma 4, it follows that if $\bar{\Delta} = \bar{v}_{r+1} - \bar{v}_r$, then $\bar{\Delta} = (I - Q_{r+1})^{-1}\bar{\delta}$, where $\bar{\delta} = (Q_{r+1}\bar{v}_r + \bar{b}_{r+1}) - (Q_r\bar{v}_r + \bar{b}_r)$. Also $(I - Q_{r+1})^{-1}$ has positive entries along the diagonal and all entries are non-negative. So it remains to show that $\bar{\delta} \geq 0$ and that some entry is > 0.

As in Lemma 4, it can be shown that the ith entry of $\bar{\delta}$ is > 0 and that if $l \neq i'$ and $l \neq i$, the lth entry of $\bar{\delta}$ is 0. It remains to consider $\bar{\delta}(i')$ in the case that $i' \leq t$. Since $\tau(r)$ is optimal with respect to $\sigma(r)$ and $(i',j') \in \tau(r)$, it must be the case that $v_r(j') \leq v_r(k')$. But the i'th entry of $Q_{r+1}\bar{v}_r + \bar{b}_{r+1}$ and $Q_r\bar{v}_r + \bar{b}_r$ are $v_r(k')$ and $v_r(j')$ respectively. Hence the i'th entry of $\bar{\delta}$ is $v_r(k') - v_r(j') \geq 0$. This concludes the argument for the vertices $1,\ldots,t$.

Next we show that if $v_{r+1}(l) = 0$ then also $v_r(l) = 0$. It is equivalent to show that all vertices with a path to the 1-sink vertex in $G_r = G_{\sigma(r),\tau(r)}$ also have a path to the 1-sink in $G_{r+1} = G_{\sigma(r+1),\tau(r+1)}$. We first show that replacing (i,j) by (i,k) in G_r does not disconnect any vertex from the 1-sink. This is true because $v_r(j) < v_r(k)$, since (i,k) replaces (i,j) in the construction of σ_{r+1}. Hence k must be connected to the 1-sink and so i is. Therefore, $v_r(l) > 0 \Rightarrow v_{\sigma(r+1),\tau(r)}(l) > 0$.

It remains to argue that by further replacing (i',j') by (i',k') in $G_{\sigma(r+1),\tau(r)}$ to obtain G_{r+1}, any vertex in G_r connected to the 1-sink is also connected to the 1-sink in G_{r+1}. This is immediate when $v_r(i') = 0$, since in this case i' is not connected to the 1-sink in G_r. When $v_r(i') > 0$, it must be the case that $v_r(j') > 0$, since i' is connected to j' in G_r. Also $v_r(j') \leq v_r(k')$ since $\tau(r)$ is the optimal strategy with respect to $\sigma(r)$. Therefore, $v_r(k') > 0$. Since we have already shown that replacing (i,j) by (i,k) to obtain $G_{\sigma(r+1),\tau(r)}$ does not disconnect any vertex from the 1-sink, it must also be the case that $v_{\sigma(r+1),\tau(r)}(k') > 0$. Hence replacing (i',j') by (i',k') in $G_{\sigma(r+1),\tau(r)}$ to obtain G_{r+1} cannot disconnect any vertex of G_r from the 1-sink, as required. □

Recall that we defined the value of a *SSG* G to be the $\max_\sigma \min_\tau v_{\sigma,\tau}(start)$.

An interesting question is whether it is equivalent to switch the order of max and min quantifiers in this definition. In the next lemma, we show that both the max-min and the min-max definitions are equivalent. The proof is a straightforward application of the techniques of the preceding lemmas. Similar results have been proved for many types of stochastic games (see Peters and Vrieze [8]), but this result is stronger than others because of our restriction to pure strategies of the players.

Lemma 6 *For any simple stochastic game, and any vertex l of the game,*

$$\min_\tau \max_\sigma v_{\sigma,\tau}(l) = \max_\sigma \min_\tau v_{\sigma,\tau}(l).$$

Proof: It is easy to show that $\max_\sigma \min_\tau v_{\sigma,\tau}(l) \le \min_\tau \max_\sigma v_{\sigma,\tau}(l)$. To prove the other direction, we show that

$$\min_\tau \max_\sigma v_{\sigma,\tau}(l) \le \max_\sigma v_{\sigma,\tau'}(l) \le v_{\sigma',\tau'}(l) \le \min_\tau v_{\sigma',\tau}(l) \le \max_\sigma \min_\tau v_{\sigma,\tau}(l),$$

where σ' and τ' are the optimal strategies of Lemmas 4 and 5. The first and last inequality are straightforward, so we only consider the other two. We first show that $v_{\sigma',\tau'}(l) \le \min_\tau v_{\sigma',\tau}(l)$. That is, for any fixed strategy τ, $v_{\sigma',\tau'}(l) \le v_{\sigma',\tau}(l)$. Without loss of generality suppose that for $l < n$, $v_{\sigma',\tau}(l) > 0$ if and only if $l \in \{1, \ldots, t\}$. Let $\bar{v}_{\sigma',\tau} = (v_{\sigma',\tau}(1), \ldots, v_{\sigma',\tau}(t))^T$ and $\bar{v}_{\sigma',\tau'} = (v_{\sigma',\tau'}(1), \ldots, v_{\sigma',\tau'}(t))^T$. Let $\bar{v}_{\sigma',\tau} = Q\bar{v}_{\sigma',\tau} + \bar{b}$, where Q and $\bar{b}$ satisfy the conditions of Lemma 1.

Then we claim that $\bar{v}_{\sigma',\tau'} \le Q\bar{v}_{\sigma',\tau'} + \bar{b}$. This is because if i is not a min vertex, the ith component of both sides of the inequality are equal; if i is a min vertex, the value of the left hand side is the minimum of the value of its neighbors whereas the value of the right hand side is one of these neighbors. Finally, since $I - Q$ is invertible and all entries of $(I - Q)^{-1}$ are ≥ 0, $\bar{v}_{\sigma',\tau'} \le (I - Q)^{-1}\bar{b} = \bar{v}_{\sigma',\tau}$. Thus for $1 \le l \le t$, $v_{\sigma',\tau'}(l) \le \bar{v}_{\sigma',\tau}(l)$.

Next consider l such that $v_{\sigma',\tau}(l) = 0$. We must show that for all such l, $v_{\sigma',\tau'}(l) = 0$. It is sufficient to show that all vertices with a path to the 1-sink in $G_{\sigma',\tau'}$ also have a path to the 1-sink in $G_{\sigma',\tau}$. We consider the special case where τ' and τ differ in only one edge, say $(i,j) \in \tau'$ and $(i,k) \in \tau$. If $v_{\sigma',\tau'}(i) = 0$ then replacing (i,j) by (i,k) cannot disconnect any vertex in the graph from the 1-sink. It remains to consider the case where $v_{\sigma',\tau'}(i) > 0$. But then j is connected to the 1-sink and since $v_{\sigma',\tau'}(k) \ge v_{\sigma',\tau'}(j)$, k must also be connected to the 1-sink. Hence again replacing (i,j) by (i,k) does not disconnect any vertices from the 1-sink.

Finally, the proof that $\max_\sigma v_{\sigma,\tau'}(l) \le v_{\sigma',\tau'}(l)$ is very similar to the proof just given. □

Let $\bar{v}_{opt}$ be the n-vector whose ith component is $\max_\sigma \min_\tau v_{\sigma,\tau}(l)$. We call $\bar{v}_{opt}$ the *value vector* of G. Then $\bar{v}_{opt} = \bar{v}_{\sigma',\tau'}$ where σ' and τ' are the optimal strategies defined in Lemmas 4 and 5 and so $\bar{v}_{opt}(start)$ is the value of the game.

We complete this section with one observation about the vector $\bar{v}_{opt}$. For any *SSG* G, let I_G be the function $I_G : [0,1]^n \to [0,1]^n$, (written simply as I if there is no ambiguity about which graph is meant), defined as follows. $I_G(\bar{x}) = \bar{y}$ where

$$y(i) = \begin{cases} \max\{x(j), x(k)\}, & \text{if } i \text{ is a max vertex of } G \text{ with neighbors } j,\ k, \\ \min\{x(j), x(k)\}, & \text{if } i \text{ is a min vertex of } G \text{ with neighbors } j,\ k, \\ \frac{1}{2}(x(j) + x(k)), & \text{if } i \text{ is an average vertex of } G \text{ with neighbors } j,\ k, \\ 0, & \text{if } i = n - 1, \\ 1, & \text{if } i = n. \end{cases}$$

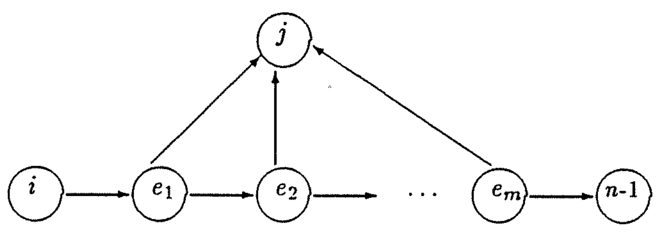
We call the set of vectors $\bar{z}$ for which $\bar{z} = I(\bar{z})$ the *solutions* of G. Then from Lemma 5, it is straightforward to show that $\bar{v}_{opt}$ is a solution of G.

3 Complexity Results

In this section we show that the *SSG* value problem is in *NP* $\cap$ co-*NP*. A property of any *SSG* G that we will use is that the value vector $\bar{v}_{opt}$ of G is a solution of G; that is, $I(\bar{v}_{opt}) = \bar{v}_{opt}$. Moreover, given a vector $\bar{z}$ is is easy to check whether it is a solution of G or not. These facts form the basis of the proof that the *SSG* problem is in *NP* $\cap$ co-*NP*. Roughly, if a game G has a ***unique*** solution, then a nondeterministic algorithm can guess the solution and it can be deterministically checked whether the value of the start vertex in that solution is greater than 1/2 or not. In general however, the solution of a simple stochastic game may not be unique.

Therefore, we show that in time polynomial in the size of an *SSG* G, we can construct from G a new game G' such that G' has a unique solution and the value of $G' > 1/2$ if and only if the value of $G > 1/2$. The new game G' is called a *stopping SSG* because for any pair of strategies σ, τ of the players in the new game, there is a path from every vertex of $G_{\sigma,\tau}$ to a sink vertex. This ensures that the game ends with probability 1 and this fact will be used to prove that G' has a unique solution. (The term "stopping stochastic game" was introduced by Shapley [10]. Stopping games are also known as *discounted* stochastic games).

For any integer $m > 0$, a stopping *SSG* G' is obtained from an arbitrary *SSG* G as follows. G' contains all the vertices $\{1, 2, \ldots, n\}$ of G (where as usual, $n-1$ and n are the 0- and 1-sink vertices) and in addition, for every edge e of G, G' contains a set of average vertices $\{e_1, e_2, \ldots, e_m\}$. None of the original edges of G are in G'; instead, for each edge $e = (i, j)$ of G, the following set of edges is included in G': $\{(i, e_1), (e_1, e_2), (e_2, e_3), \ldots, (e_{m-1}, e_m), (e_m, n-1), (e_1, j), (e_2, j), \ldots, (e_m, j)\}$.



From this construction, note that if a path is followed from vertex i, the first vertex from the set $\{1, \ldots, n\}$ that is reached is either vertex j or vertex $n-1$ (the 0-sink); and this vertex is reached in at most $m+1$ steps. Therefore, with probability $1/2^m$ the 0-sink is the first vertex from the set $\{1, \ldots, n\}$ that is reached on a random walk from i. If the 0-sink vertex is reached, the game ends. The resulting stopping game is called a $1/2^m$-stopping game because at any step of the game, from any vertex in the set $\{1, \ldots, n\}$, the probability of ending the game before reaching another vertex in the set $\{1, \ldots, n-2\}$ is at least $1/2^m$.

Shapley [10] showed that any stopping game constructed in this manner has a unique solution; we include the proof here for completeness.

Lemma 7 (Shapley, [10]) *If G' is a $1/2^m$-stopping game corresponding to G then G' has a unique solution.*

Proof: Suppose that G has n vertices. By definition, for any solution $\bar{z}$ of G, $\bar{z} = I(\bar{z})$. Assume that the vertices numbered $1, \ldots, n$ in G' are the vertices that are also in G. First, note that the values of the new average vertices of G', that is, the vertices not in $\{1, \ldots, n\}$, are uniquely determined by the values of the vertices in the set $\{1, \ldots, n\}$. Hence if $\bar{w}$ and $\bar{x}$ are distinct solutions of G', they must differ in at least one of their first n components, i.e. $w(i) \neq x(i)$ for some $i \leq n$. Also if $\bar{z}'$ is a solution of G' then for $m > 0$ and $1 \leq i \leq n-2$,

$$\bar{z}'(i) = (1 - 1/2^m)(I_G(\bar{z}'))(i).$$

This can be proven easily by induction on m. Intuitively it is true for value vectors z' of G' because if i has children j and k in G, then in the game G', with probability $1/2^m$ the game ends before the token is moved from i to j or k.

Now suppose to the contrary that the vectors $\bar{w}$ and $\bar{x}$ are two distinct solutions of G'. We show that this leads to a contradiction. In the rest of this proof, let $||\bar{w} - \bar{x}|| = \max_{1 \leq i \leq n} |w(i) - x(i)| = c > 0$. Then $||\bar{w} - \bar{x}|| = (1 - 1/2^m)||I(\bar{w}) - I(\bar{x})||$. Moreover, it is straightforward to show that $||I(\bar{w}) - I(\bar{x})|| \leq c$, since for each $i, 1 \leq i \leq n-2$ with neighbors j and k, if $I(\bar{w})(i)$ denotes the ith component of $I(\bar{w})$ then

$$|(I(\bar{w}))(i) - (I(\bar{x}))(i)| = \begin{cases} |\max(w(j), w(k)) - \max(x(j), x(k))|, & \text{if } i \in V_{max}, \\ |\min(w(j), w(k)) - \min(x(j), x(k))|, & \text{if } i \in V_{min}, \\ 1/2(w(j) + w(k)) - 1/2(x(j) + x(k)), & \text{if } i \in V_{average}. \end{cases}$$

If $|w(i) - x(i)| \leq c$ for $1 \leq i \leq n-2$ then each of these differences is also $\leq c$.

We have now shown that

$$c \leq ||\bar{w} - \bar{x}|| \leq (1 - 1/2^m)||I(\bar{w}) - I(\bar{x})|| \leq (1 - 1/2^m)c,$$

a contradiction. Hence the stopping game G' has a unique solution. □

The last step to proving the main theorem is completed in the next lemma, where it is shown that the value of a *SSG* is $> 1/2$ if and only if the value of the corresponding $(1/2^{cn})$-stopping game is $> 1/2$, for some constant $c > 0$. This lemma is a discrete version of a result of Filar [3] who proved that the value of G is the limit, as $\beta \to 0$, of the value of the β-stopping game corresponding to G.

Lemma 8 *There is a constant $c > 0$, such that if G is a* SSG *with n vertices, the value of G is $> 1/2$ if and only if the value of the corresponding $(1/2^{cn})$-stopping game is $> 1/2$.*

Proof: One direction of this proof is straightforward; it is not hard to see that the stopping *SSG* corresponding to any *SSG* must have lower value. Hence if the value of $G \leq 1/2$ it must be the case that the value of any β-stopping game is $\leq 1/2$.

Let G' be the β-stopping *SSG* corresponding to G for some $\beta = 1/2^{cn}$ and let $\bar{v}'$ be the value vector of G'. Since G and G' have the same number of max and min vertices, there is a 1-1 correspondence between the strategies of both games. We show that for any pair of strategies σ and τ of the players, for any $i \in \{1, \ldots, n\}$, the values of i with respect to σ and τ in games G and G' are close. Specifically, for any vertex $i \in \{1, \ldots, n-2\}$, $v_{\sigma,\tau}(i) - v'_{\sigma,\tau}(i) \leq 1/4^n$ for sufficiently large c. Then from Lemma 2, if the value of G is greater than $1/2$ then so is the value of G'.

From the construction of G' from G, it is easy to verify that the same vertices from the set $\{1,\ldots,n-2\}$ have non-zero value in graphs $G_{\sigma,\tau}$, $G'_{\sigma,\tau}$. Let these vertices be $1,\ldots,t$, let $\bar{v}_{\sigma,\tau}=(v_{\sigma,\tau}(1),\ldots,v_{\sigma,\tau}(t))$ and $\bar{v}'_{\sigma,\tau}=(v'_{\sigma,\tau}(1),\ldots,v'_{\sigma,\tau}(t))$. Let $\bar{v}_{\sigma,\tau}=Q\bar{v}_{\sigma,\tau}+\bar{b}$ be the equation uniquely defining the values $v_{\sigma,\tau}(i), 1\le i\le t$. Again from the construction of G' it follows that $\bar{v}'_{\sigma,\tau}=(1-\beta)Q\bar{v}'_{\sigma,\tau}+(1-\beta)\bar{b}$.

Since Q is invertible, $\bar{v}_{\sigma,\tau}=(I-Q)^{-1}\bar{b}=\bar{b}\sum_{j=0}^{\infty}Q^j$. Similarly, $\bar{v}'_{\sigma,\tau}=\bar{b}\sum_{j=0}^{\infty}(1-\beta)^{j+1}Q^j$. Hence

$$\bar{v}_{\sigma,\tau}-\bar{v}_{\sigma,\tau'}=\bar{b}\sum_{j=0}^{\infty}(1-(1-\beta))^{j+1})Q^j, \text{ and so}$$

$$||\bar{v}_{\sigma,\tau}-\bar{v}_{\sigma,\tau'}||\le\sum_{j=0}^{\infty}(1-(1-\beta))^{j+1})||Q^j||.$$

Rewriting $(1-(1-\beta)^{j+1})$ as a telescoping sum, it follows that it can be bounded above by $\beta(j+1)$. To get an upper bound on $||\bar{v}_{\sigma,\tau}-\bar{v}_{\sigma,\tau'}||$, we also need a bound on the rate of convergence of $||Q^j||$ to 0 as $j\to\infty$. In the proof of Lemma 1, we showed that for all integers $m\ge 0$, $||Q^{nm}||\le(1-1/2^n)^m$. From this it follows that

$$||\bar{v}_{\sigma,\tau}-\bar{v}_{\sigma,\tau'}||\le\sum_{j=0}^{\infty}\beta(j+1)(1-1/2^n)^{\lfloor j/n\rfloor}\le\beta 2^{3n},$$

for sufficiently large n. For an appropriate choice of the constant c (say, $c=5$), if $\beta=1/2^{cn}$ then $||\bar{v}_{\sigma,\tau}-\bar{v}_{\sigma,\tau'}||<1/4^n$. By Lemma 3, it follows that if $\bar{v}_{\sigma,\tau}(start)>1/2$ then $\bar{v}_{\sigma,\tau'}(start)>1/2$.

In particular, if σ' and τ' are the optimal strategies of players 1 and 0 respectively given by Lemma 5, $\bar{v}_{\sigma',\tau'}(start)>1/2$ implies that $\bar{v}'_{\sigma',\tau'}(start)>1/2$. Therefore, the value of G is $>1/2$ if and only if the value of G' is $>1/2$. □

Theorem 1 *The* SSG *value problem is in* NP $\cap$ co-NP.

Proof: We first describe a nondeterministic Turing machine M that on input a *SSG* G with n vertices, accepts if and only if G has value $>1/2$. M constructs the $1/2^{cn}$-stopping game G' corresponding to G, where c is the constant of Lemma 8 and then guesses a vector $\bar{z}\in[0,1]^n$, where each component of $\bar{z}$ is rational. It then verifies that the vector is a solution of G'. If so and if the value of the start vertex is $>1/2$ then it accepts; else it rejects. From Lemma 8, M accepts G if and only if the value of G is $>1/2$.

A similar construction shows that the complement of the *SSG* value problem is in *NP*. We construct a machine $\bar{M}$ that accepts a *SSG* G if and only if its value is $\le 1/2$. On input G, $\bar{M}$ guesses a vector $\bar{z}$ as before; verifies that the vector is a solution of G and that the value of the start vertex is $\le 1/2$. If so, $\bar{M}$ accepts; else it rejects. □

4 Extensions and Open Problems

There are many ways to generalize simple stochastic games. We describe some of these generalizations in this section and consider whether the results of the previous sections extend to these generalized classes of stochastic games.

A natural generalization is to allow more than two edges from a vertex and to allow edges from average vertices to be labeled with arbitrary rational probabilities, such that the probabilities from each average vertex sum to 1. Suppose G is a stochastic game with n vertices generalized in this way such that all edges from average vertices of G are labeled with probabilities of the form p/q where $0 \le p \le q \le 2^m$. Then, as we outline next, in time polynomial in $\max\{m, n\}$, G can be transformed to a *SSG* G' such that the value of G equals the value of G'. Hence the value problem for these stochastic games is also in $NP \cap$ co-NP.

Briefly, G' can be constructed from G in two stages. In the first stage, the out degree of every vertex is reduced to two by replacing each node with fanout l where $l > 2$ by a binary tree with l leaves, and relabeling the edges with the appropriate probabilities. In the second stage, if i is a vertex with two children j and k and edges labeled with probabilities p_1/q, p_2/q, where $2^t < q \le 2^{t+1}$, then i is replaced by a complete binary tree of average vertices of depth $t+1$, rooted at i. All outgoing edges of p_1 of the leaves of this tree point to j; all outgoing edges of p_2 of the leaves of this tree point to t; and the remaining leaves point to i, the root of the tree. This construction guarantees that the probability of reaching j and k from i is exactly p_1/q and p_2/q, respectively.

Another generalization, standard in the literature on stochastic games is to allow *payoffs* at the vertices. Simple stochastic games with payoffs were studied by Filar [3]. Suppose that each vertex of a stochastic game is labeled with a rational number, called its payoff. Let the payoff associated with vertex i be denoted by a_i (the payoffs can be negative). Assume that the payoffs are of the form p/q where $1 \le p, q \le 2^n$. The interpretation of the payoff in the game is that whenever the token is placed on vertex i, player 0 must pay amount a_i to player 1. In such games, the value of a game may not be well-defined in general. Hence we only consider stopping games, where the expected total winnings of either player on any play of the game is finite. That is, we assume that there is one special sink vertex with payoff 0 and that for all pairs of strategies of the players, this vertex is eventually reached with probability 1 on any play of the game.

Let G be a simple stochastic game with payoffs, such that G is a stopping game. For a fixed pair of strategies σ and τ of the players, the value of the vertices of G with respect to σ and τ is the unique solution to the equation $\bar{v}_{\sigma,\tau} = Q\bar{v}_{\sigma,\tau} + \bar{a}$, where $\bar{a}$ is the vector of payoffs of the vertices and Q is the 1-step transition matrix of the graph $G_{\sigma,\tau}$, excluding the special sink vertex. Lemmas 4 – 6 of Section 2.1 generalize to show that there is a pair of optimal pure strategies σ', τ' of players 1 and 0 respectively such that for all vertices l,

$$\min_{\tau} \max_{\sigma} v_{\sigma,\tau}(l) = v_{\sigma',\tau'}(l) = \max_{\sigma} \min_{\tau} v_{\sigma,\tau}(l).$$

Also the definitions of an iteration function and a solution, given in Section 3, generalize to *SSG*'s with payoffs. Thus the results of that section can be extended to show that the value problem for stopping *SSG*'s with payoffs is in $NP \cap$ co-NP.

The last generalization we consider was introduced by Shapley [10] in his original paper on stochastic games. In this model, the players move *simultaneously* at each step of the game. The vertices of the graph are partitioned into two groups, which we call the ***average*** and the ***strategic*** vertices. We assume that there are four edges from each strategic vertex, each labeled by a tuple (b_0, b_1), where $b_0, b_1 \in \{0, 1\}$. In playing the game, when the token is placed on a strategic vertex, players 0 and 1 simultaneously choose a binary value for bits b_0 and b_1, respectively. The pair of bits chosen determines to which neighbor the token is moved.

Because of the simultaneity in this model, in order for the minimax result (Lemma 6) to hold, it is necessary to generalize the definition of a strategy of a player. A *mixed* (or probabilistic) strategy of player 0 is a set of pairs (p_{0i}, p_{1i}), one per strategic vertex i, such that $p_{0i} + p_{1i} = 1$. The interpretation of this strategy is that when the token is on vertex i, player 0 chooses value 0 for b_0 with probability

p_{0i} and value 1 for b_0 with probability p_{1i}. A mixed strategy of player 1 is defined similarly. With this definition of strategy, the results of Section 2 extend to simultaneous stochastic games (see Shapley [10]). However, it is an open problem whether the value problem for simultaneous stochastic games is in *NP* $\cap$ co-*NP*. One obstacle to solving this is that the value of a simultaneous stochastic game need not be rational - see [8] for an example of this.

Finally, a challenging open problem is whether the *SSG* value is in *P*. A related problem which might be of significance in solving this is the following. Is there a polynomial time algorithm that separates graphs with high value from those with low value? More precisely, let A be the set of *SSG*'s that have value $> 3/4$ and let B be the set of *SSG*'s that have value $< 1/4$. Then it would be of interest to find a polynomial time decision algorithm M that, on input $G \in A \cup B$, accepts if and only if $G \in A$.

5 Acknowledgement

I thank Lazlo Babai and Mario Szegedy for some enlightening discussions on simple stochastic games. They pointed out the interesting combinatorial structure underlying a complexity class I was investigating and provided many good insights into its properties.

References

[1] A. K. CHANDRA, D. C. KOZEN, and L. J. STOCKMEYER. Alternation. *Journal of the Association of Computing Machinery*, 28(1):114–133, 1981.

[2] C. DERMAN. *Finite State Markov Decision Processes.* Academic Press, 1972.

[3] J. A. FILAR. Ordered field property for stochastic games when the player who controls transitions changes from state to state. *Journal of Algorithms*, 34:503–515, 1981.

[4] J. GILL. The computational complexity of probabilistic turing machines. *SIAM Journal on Computing.*, 6:675–695, 1977.

[5] HOWARD. *Dynamic Programming and Markov Processes.* M.I.T. Press, 1960.

[6] J. G. KEMENY and J. LAURIE SNELL. *Finite Markov Chains.* D. Van Nostrand Company, Inc., 1960.

[7] L. G. KHACHIYAN. A polynomial algorithm in linear programming. *Soviet Math Dokl.*, 20:191–194, 1979.

[8] H. J. M. PETERS and O. J. VRIEZE. *Surveys in game theory and related topics, CWI Tract 39.* Centrum voor Wiskunde en Informatica, Amsterdam, 1987.

[9] W. L. RUZZO, J. SIMON, and M. TOMPA. Space-bounded hierarchies and probabilistic computations. *Journal on Computer and System Sciences*, 28(2):216–230, April 1984.

[10] L. S. SHAPLEY. Stochastic games. In *Proceedings of the National Academy of Sciences, U.S.A*, volume 39, pages 1095–1100, 1953.

[11] O.J. VRIEZE. *Stochastic games with finite state and action spaces, CWI Tract 33.* Centrum voor Wiskunde en Informatica, Amsterdam, 1987.