

# The Complexity of Stochastic Games

ANNE CONDON\*

*Computer Sciences Department, University of Wisconsin–Madison,  
1210 West Dayton Street, Madison, Wisconsin 53706*

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  $NP \cap co-NP$ . © 1992 Academic Press, Inc.

## 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 are 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  $\frac{1}{2}$ .

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 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:

\* Supported by National Science Foundation Grant DCR-8802736.

- (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  $\frac{1}{2}$ —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  $NP \cap \text{co-}NP$ . Although many number theoretic problems not known to be in  $P$  lie in the class  $NP \cap \text{co-}NP$ , combinatorial problems that lie between  $P$  and  $NP \cap \text{co-}NP$  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 (1953) and many variations of the model have been investigated since then (see Peters and Vrieze, 1987, for a survey). The class of stochastic games studied in this paper is somewhat similar to the class of stochastic games with switching transitions, studied by Filar (1981). 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 Chandra *et al.*, 1981.) A polynomial time algorithm for the simple stochastic game problem would 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  $NP \cap \text{co-}NP$ . We also describe some special cases where the problem can be solved in polynomial time, namely when the graph is restricted to have just two of the three types of vertices max, min, and average. 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  $V_{\max}$ ,  $V_{\min}$ ,  $V_{\text{average}}$ , 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, ...,  $n$ , with  $n-1$  and  $n$  being the 0- and 1-sink vertices, respectively. An example of a SSG is given in Fig. 1.

Associated with the game are two players, 0 and 1. A *strategy*  $\tau$  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  $\tau$ . Informally, if  $(i, j) \in \tau$  then in a game where player 0 uses strategy  $\tau$ , the token is always moved from vertex  $i$  to vertex  $j$ . Similarly, a strategy  $\sigma$  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  $\sigma$ . 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, 1987.)

Corresponding to strategy  $\sigma$  is a graph  $G_\sigma$ , which is the subgraph of  $G$  obtained by removing from each max vertex the outgoing edge that is not in the strategy  $\sigma$ . Similarly, corresponding to a pair of strategies  $\sigma$  and  $\tau$ , is a graph  $G_{\sigma, \tau}$  obtained from  $G_\sigma$  by removing from each min vertex the

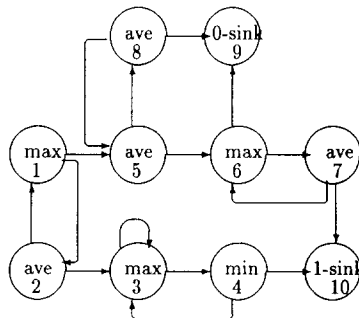


FIG. 1. A simple stochastic game with 10 vertices. Vertex 1 is the start vertex.

outgoing edge that is not in  $\tau$ . 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  $p_{nn} = p_{n-1, n-1} = 1$ ,  $p_{nj} = 0$  if  $j \neq n$  and  $p_{n-1, j} = 0$  if  $j \neq n-1$ . We say a SSG *halts with probability 1* if for all pairs of strategies  $\sigma, \tau$ , every vertex in  $G_{\sigma, \tau}$  has a path to a sink vertex.

We define the *value*  $v_{\sigma, \tau}(i)$  of each vertex  $i$  of  $G$  with respect to strategies  $\sigma$  and  $\tau$  to be the probability that player 1 wins the game if the start vertex is  $i$  and the players use strategies  $\sigma$  and  $\tau$ . 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 Fig. 2, the subgraph of the SSG of Fig. 1 is given, for strategies  $\sigma = \{(1, 5), (3, 4), (6, 7)\}$  and  $\tau = \{(4, 3)\}$  of player 1 and 0, respectively.

Without loss of generality, assume that the vertices of  $G_{\sigma, \tau}$  are numbered so that  $\{1, \dots, t\}$  is the set of vertices from which a sink vertex is reachable, other than the sink vertices themselves. 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

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

We next show that the equations (1) have a unique solution.

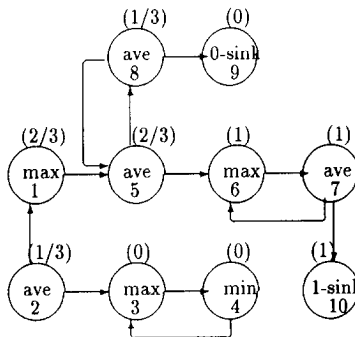


FIG. 2. Subgraph of the SSG of Fig. 1, for strategies  $\sigma = \{(1, 5), (3, 4), (6, 7)\}$  and  $\tau = \{(4, 3)\}$  of player 1 and 0, respectively. The value  $v_{\sigma, \tau}(i)$  of each vertex  $i$  is given above the vertex.

LEMMA 1. Let  $G$  be a SSG with  $n$  vertices and let  $\sigma$  and  $\tau$  be strategies of players 1 and 0, respectively. Suppose that the vertices of  $G_{\sigma, \tau}$  are numbered so that the set of vertices (other than the sink vertices) from which a sink vertex is reachable is exactly the set  $\{1, \dots, t\}$  and let  $\bar{v}_{\sigma, \tau} = (v_{\sigma, \tau}(1), \dots, v_{\sigma, \tau}(t))$ . Then there are an  $t \times t$  matrix  $Q$  and a  $t$ -vector  $\bar{b}$  with entries in  $\{0, \frac{1}{2}, 1\}$  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 Eqs. (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 1-step transition matrix of the vertices  $\{1, \dots, t\}$ . For each  $i$ , the  $i$ th component of  $\bar{b}$  is from the set  $\{0, \frac{1}{2}, 1\}$  and equals the probability of reaching the 1-sink vertex from  $i$  in 1 step. Also all entries of  $Q$  are in  $\{0, \frac{1}{2}, 1\}$ . 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 show that  $\lim_{l \rightarrow \infty} Q^l = 0$ . 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  $l \rightarrow \infty$ ,  $Q^l \rightarrow 0$ . Note that the  $ij$ th 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  $i$ th row of  $Q^{mn}$  is at most 1 minus the probability of reaching a sink vertex of  $G_{\sigma, \tau}$  from  $i$  in  $n$  steps. Since a sink vertex is reachable from  $i$ , there must be a path of length  $\leq n$  in  $G_{\sigma, \tau}$  from  $i$  to a sink. Hence the probability of reaching a 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.

Since  $\lim_{l \rightarrow \infty} Q^l = 0$ , 1 cannot be an eigenvalue of  $Q$ . Hence  $I - Q$  is invertible. In fact, since  $I - Q^l = (I - Q)(I + Q + Q^2 + \dots + Q^{l-1})$ , multiplying both sides by  $(I - Q)^{-1}$  we get that

$$(I - Q)^{-1} (I - Q^l) = I + Q + Q^2 + \dots + Q^{l-1}.$$

Again, since  $\lim_{l \rightarrow \infty} Q^l = 0$ , the limit of the left hand side of this equation is  $(I - Q)^{-1}$ ; hence the limit of the right hand side must exist and  $(I - Q)^{-1} = I + Q + Q^2 + \dots$ . 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

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

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  $\sigma$  and  $\tau$  such that the value of the game equals  $v_{\sigma, \tau}(\text{start})$ . To investigate the complexity of these problems, we consider the following decision problem for SSGs.

The *SSG value problem* is: Given a SSG, is its value  $> \frac{1}{2}$ ?

In Section 3, we show that this decision problem is in the class  $NP \cap \text{co-}NP$ . 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  $\frac{1}{2}$  if it is not equal to  $\frac{1}{2}$ . 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,  $0 \leq p, q \leq 4^{n-1}$ .*

*Proof.* The proof of this lemma is very similar to a proof by Gill (1977, Lemma 6.6). Let  $\sigma$  and  $\tau$  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  $\sigma$  and  $\tau$  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 a path to a sink vertex, excluding the sink vertices, are numbered from 1 to  $t$  and let  $\bar{v}_{\sigma, \tau} = (v_{\sigma, \tau}(1), \dots, v_{\sigma, \tau}(t))$ . 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  $i \in \{1, 2, \dots, t\}$  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  $i$ th column of  $2(I - Q)$  by  $2\bar{b}$ . The determinant of  $2(I - Q)$  is at most  $4^t$ . To prove this, we use the following property of  $2(I - Q)$ :

Each row of  $2(I - Q)$  has a constant number of non-zero entries from the set  $\{0, \pm 1, \pm 2\}$ , whose absolute values sum to at most 4. (\*)

We show by induction on  $t$  that any  $t \times t$  matrix  $A$  with property (\*) has determinant at most  $2^t$ . If  $t = 1$ , the determinant is at most 2. If  $t > 1$ , then by the induction hypothesis, the absolute value of the determinant of any minor of  $A$  is at most  $4^{t-1}$ . This is because the minors of  $A$  also have property (\*). Hence the determinant of  $A$  is at most  $4^{t-1}$  times the sum of the absolute values of the entries in the first row of  $A$ . This product is at most  $4^t$ , again applying property (\*).

Thus  $D$  is at most  $4^t \leq 4^{n-1}$ , since  $t < n$ . 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  $4^{n-1}$ . In particular this is true for the start vertex of  $G_{\sigma, \tau}$ .

Since the strategies  $\sigma$  and  $\tau$  were chosen arbitrarily,  $\max_{\sigma} \min_{\tau} v_{\sigma, \tau}(\text{start})$  must be the quotient of two numbers, each of value at most  $4^{n-1}$ . ■

**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.* From Lemma 2, the value of a simple stochastic game with  $n$  vertices is of the form  $p/q$  where  $p$  and  $q$  are integers,  $0 \leq p, q \leq 4^{n-1}$ . If  $p/q > 1/2$ , then  $2p \geq q + 1$ . Hence,

$$p/q = 2p/2q \geq (q + 1)/2q = 1/2 + 1/2q \geq 1/2 + (1/2) 4^{n-1} > 1/2 + 1/4^n. \quad \blacksquare$$

In the following lemmas, we describe properties of optimal strategies of the players of a SSG. For simplicity, we only consider SSGs that halt with probability 1 in the rest of this section. (Recall that a SSG halts with probability 1 if on all pairs of strategies  $\sigma, \tau$ , a sink vertex is reachable from every vertex of  $G_{\sigma, \tau}$ .) We first show that for each strategy  $\sigma$  of player 1, there is a strategy  $\tau(\sigma)$  of player 0 that is "optimal" with respect to  $\sigma$  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  $\tau(\sigma)$  an optimal strategy of player 0 with respect to strategy  $\sigma$ .

**LEMMA 4** (Howard, 1960). *Let  $G$  be a simple stochastic game with  $n$  vertices that halts with probability 1 and let  $\sigma$  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 (1960) described an algorithm to construct  $\tau(\sigma)$  satisfying the lemma. The algorithm proceeds in iterations. There is a current strategy  $\tau(q)$  for each iteration  $q$ ; the current strategy of the initial iteration is chosen arbitrarily and is denoted by  $\tau(0)$ . Strategy  $\tau(q+1)$  is constructed from strategy  $\tau(q)$  by finding any one vertex  $i \in V_{\min}$  with neighbors  $j, k$  such that  $(i, j) \in \tau(q)$  but  $v_{\sigma, \tau(q)}(k) < v_{\sigma, \tau(q)}(j)$  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  $v_q(i) = v_{\sigma, \tau(q)}(i)$ . The fact that Howard's algorithm halts follows from the following property, which we will prove: if the algorithm does not halt at round  $q$  then strategy  $\tau(q+1)$  improves strategy  $\tau(q)$  in that for each vertex  $l$  of  $G$ ,  $v_{q+1}(l) \leq v_q(l)$  and for at least one vertex the inequality is strict. This ensures that no strategy can be repeated in the sequence  $\tau(0), \tau(1), \dots, \tau(\sigma)$ . Since there are only a finite number of strategies (at most  $2^n$ ) for player 1, the sequence is finite.

Clearly, for the 0- and 1-sink vertices, which are numbered  $n-1$  and  $n$ ,  $0 = v_{q+1}(n-1) = v_q(n-1)$  and  $1 = v_{q+1}(n) = v_q(n)$ . Therefore we restrict our attention to the vertices numbered  $1, \dots, n-2$ . Let  $t = n-2$ . Let  $\bar{v}_q < (v_q(1), \dots, v_q(t))^T$ . As in Lemma 1 let matrices  $Q_q, Q_{q+1}$  and vectors  $\bar{b}_q$  and  $\bar{b}_{q+1}$  be such that  $\bar{v}_q = Q_q \bar{v}_q + \bar{b}_q$  and  $\bar{v}_{q+1} = Q_{q+1} \bar{v}_{q+1} + \bar{b}_{q+1}$ . Let  $\bar{A} = \bar{v}_q - \bar{v}_{q+1}$ . We show that  $\bar{A} \geq 0$  and that some entry is actually  $> 0$ . Adding and subtracting  $Q_{q+1} \bar{v}_q + \bar{b}_{q+1}$  to  $\bar{A}$  we see that

$$\begin{aligned} \bar{A} &= (Q_q \bar{v}_q + \bar{b}_q) - (Q_{q+1} \bar{v}_q + \bar{b}_{q+1}) + (Q_{q+1} \bar{v}_q + \bar{b}_{q+1}) \\ &\quad - (Q_{q+1} \bar{v}_{q+1} + \bar{b}_{q+1}). \end{aligned}$$

If  $\delta = (Q_q \bar{v}_q + \bar{b}_q) - (Q_{q+1} \bar{v}_q + \bar{b}_{q+1})$ , then  $\bar{A} = Q_{q+1} \bar{A} + \delta$ . From Lemma 1 it follows that the matrix  $(I - Q_{q+1})$  is invertible, since  $G$  halts with probability 1. Hence  $\bar{A} = (I - Q_{q+1})^{-1} \delta$ . Also from Lemma 1, all the entries of  $(I - Q_{q+1})^{-1}$  are  $\geq 0$  and the entries along the diagonal are  $> 0$ . Therefore, it remains to show that  $\delta \geq 0$  and that one entry is actually  $> 0$ . Suppose edge  $(i, k)$  replaces  $(i, j)$  in the construction of  $\tau(q+1)$  from  $\tau(q)$ . Hence  $Q_q$  and  $Q_{q+1}$  differ in at most the  $i$ th row and similarly  $\bar{b}_q$  and  $\bar{b}_{q+1}$  differ in at most the  $i$ th row. Thus every entry of  $\delta$  must be 0, except possibly the  $i$ th entry.

Moreover, the  $i$ th entries of  $Q_q \bar{v}_q + \bar{b}_q$  and  $Q_{q+1} \bar{v}_q + \bar{b}_{q+1}$  are  $v_q(j)$  and  $v_q(k)$ , respectively; hence the  $i$ th entry of  $\delta$  is  $v_q(j) - v_q(k)$ . Since edge  $(i, k)$  replaces  $(i, j)$  in the construction of  $\tau(q+1)$  from  $\tau(q)$ , it must be the case



that  $v_q(j) > v_q(k)$  and hence the  $i$ th entry of  $\delta$  must be  $> 0$ . This completes the proof that Howard's algorithm halts. ■

The next lemma shows that in any SSG  $G$  that halts with probability 1, player 1 has a strategy  $\sigma'$  which is optimal in the following sense. Suppose player 0 uses some optimal strategy  $\tau(\sigma')$  with respect to strategy  $\sigma'$ . Then for each vertex  $i \in V_{\max}$ , the value of  $i$  with respect to strategies  $\sigma'$  and  $\tau(\sigma')$  is the maximum of the values of its children. We call any such pair of strategies  $\sigma'$ ,  $\tau' = \tau(\sigma')$  an optimal pair of strategies. The proof of this lemma is a straightforward extension of Lemma 4; similar proofs for other classes of stochastic games are surveyed in (Peters and Vrieze, 1987).

**LEMMA 5.** *Let  $G$  be a simple stochastic game with  $n$  vertices that halts with probability 1. Then there is a strategy  $\sigma'$  of player 1 such that, for some optimal strategy  $\tau' = \tau(\sigma')$  of player 0 with respect to strategy  $\sigma'$ , 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  $\sigma'$  can be found by an iterative method similar to Lemma 4. In the iterative algorithm to construct  $\sigma'$ , there is a current strategy  $\sigma(q)$  for each iteration  $q$ ; the current strategy for the initial iteration is chosen arbitrarily. Let  $\tau(q)$  be an optimal strategy of player 0 with respect to strategy  $\sigma(q)$ . Such a strategy can be found using the algorithm of Lemma 4. Let  $v_q(l)$  denote  $v_{\sigma(q), \tau(q)}(l)$ . Strategy  $\sigma(q+1)$  is constructed from strategy  $\sigma(q)$  by finding any one vertex  $i \in V_{\max}$  such that  $(i, j) \in \sigma(q)$  but  $v_q(k) > v_q(j)$  and replacing  $(i, j)$  by  $(i, k)$ . If no such vertex exists the algorithm halts; the current strategy of the final iteration is  $\sigma'$ . Let  $\tau'$  be an optimal strategy of player 0 with respect to  $\sigma'$ . 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  $v_{q+1}(l) \geq v_q(l)$  and that the inequality is strict for some  $l$ . The proof of this is complicated by the fact that at each iteration, both  $\sigma(q)$  and  $\tau(q)$  are being changed, whereas in the construction of Lemma 4 only one strategy is changed. However, the key elements in the proof are similar.

Suppose that  $(i, k)$  replaces  $(i, j)$  in the construction of  $\sigma(q+1)$  from  $\sigma(q)$  and that  $(i', k')$  replaces  $(i', j')$  in the construction of  $\tau(q+1)$  from  $\tau(q)$ . We restrict the proof to the case where  $\tau(q+1)$  and  $\tau(q)$  differ in only one edge; the proof can easily be generalized to the case when they differ in more than one edge.

Again, it is immediate that  $v_{q+1}(l) \geq v_q(l)$  for  $l = n-1$  or  $n$ . Therefore, we consider the vertices  $1, \dots, t$ , where  $t = n-2$ . Let  $\bar{v}_q = (v_q(1), \dots, v_q(t))^T$ ,  $\bar{v}_{q+1} = (v_{q+1}(1), \dots, v_{q+1}(t))^T$ ,  $\bar{v}_q = Q_q \bar{v}_q + \bar{b}_q$ , and  $\bar{v}_{q+1} = Q_{q+1} \bar{v}_{q+1} + \bar{b}_{q+1}$

for some  $Q_q$ ,  $Q_{q+1}$ ,  $\bar{b}_q$ , and  $\bar{b}_{q+1}$ . As in Lemma 4, it follows that if  $\bar{A} = \bar{v}_{q+1} - \bar{v}_q$ , then  $\bar{A} = (I - Q_{q+1})^{-1} \delta$ , where  $\delta = (Q_{q+1} \bar{v}_q + \bar{b}_{q+1}) - (Q_q \bar{v}_q + \bar{b}_q)$ . Also  $(I - Q_{q+1})^{-1}$  has positive entries along the diagonal and all entries are non-negative. So it remains to show that  $\delta \geq 0$  and that some entry is  $> 0$ .

As in Lemma 4, it can be shown that the  $i$ th entry of  $\delta$  is  $> 0$  and that if  $l \neq i'$  and  $l \neq i$ , the  $l$ th entry of  $\delta$  is 0. It remains to consider  $\delta(i')$ . Since  $\tau(q)$  is optimal with respect to  $\sigma(q)$  and  $(i', j') \in \tau(q)$ , it must be the case that  $v_q(j') \leq v_q(k')$ . But the  $i'$ th entries of  $Q_{q+1} \bar{v}_q + \bar{b}_{q+1}$  and  $Q_q \bar{v}_q + \bar{b}_q$  are  $v_q(k')$  and  $v_q(j')$ , respectively. Hence the  $i'$ th entry of  $\delta$  is  $v_q(k') - v_q(j') \geq 0$ , as required. ■

Recall that we defined the value of a SSG  $G$  to be  $\max_{\sigma} \min_{\tau} v_{\sigma, \tau}(\text{start})$ . In fact, 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, for SSGs that halt with probability 1. 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, 1987), but this result is stronger than others because of our restriction to pure strategies of the players.

LEMMA 6. *Let  $G$  be any simple stochastic game that halts with probability 1. Then for any vertex  $l$  of  $G$ ,*

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

*Proof.* It follows immediately that  $\max_{\sigma} \min_{\tau} v_{\sigma, \tau}(l) \leq \min_{\tau} \max_{\sigma} v_{\sigma, \tau}(l)$ . To prove the other direction, let  $\max_{\sigma} \min_{\tau} v_{\sigma, \tau}(l) = v_{\sigma', \tau'}(l)$ , where  $\sigma'$  and  $\tau'$  are optimal strategies. Note that we can assume that  $\sigma'$  and  $\tau'$  are optimal based on the proofs of Lemmas 4 and 5. The proof is completed once we show that

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

The first inequality is straightforward, so we show that  $\max_{\sigma} v_{\sigma, \tau'}(l) \leq v_{\sigma', \tau'}(l)$ . Fix any strategy  $\sigma$  of player 1. Clearly the statement is true if  $l$  is a sink node, so suppose  $l < n - 1$ . Let  $t = n - 2$ , let  $\bar{v}_{\sigma, \tau'} = (v_{\sigma, \tau'}(l), \dots, v_{\sigma, \tau'}(t))^T$  and  $\bar{v}_{\sigma', \tau'} = (v_{\sigma', \tau'}(1), \dots, v_{\sigma', \tau'}(t))^T$ . 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  $\bar{v}_{\sigma, \tau'} \geq Q \bar{v}_{\sigma', \tau'} + \bar{b}$ . This is because if  $i$  is not a max vertex, the  $i$ th component of both sides of the inequality are equal; if  $i$  is a max vertex, the value of the left hand side is the maximum of the values 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  $\geq 0$ ,  $\bar{v}_{\sigma', \tau'} \geq (I - Q)^{-1} \bar{b} = \bar{v}_{\sigma, \tau}$ . In particular,  $v_{\sigma', \tau'}(l) \geq \bar{v}_{\sigma, \tau}(l)$ , as required. ■

Let  $\bar{v}_{\text{opt}}$  be the  $n$ -vector whose  $i$ th component is  $\max_{\sigma} \min_{\tau} v_{\sigma, \tau}(l)$ . We call  $\bar{v}_{\text{opt}}$  the *value vector* of  $G$ . We complete this section with some observations about optimal strategies and the vector  $\bar{v}_{\text{opt}}$ . For any SSG  $G$ , let  $I_G$  be the function  $I_G: [0, 1]^n \rightarrow [0, 1]^n$  (written simply as  $I$  if there is no ambiguity about which graph is meant), defined as  $I_G(\bar{x}) = \bar{y}$ , 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 if  $G$  is a SSG that halts with probability 1, then for any pair of optimal strategies  $\sigma', \tau'$ , the vector  $\bar{v}_{\sigma', \tau'} = (v_{\sigma', \tau'}(1), \dots, v_{\sigma', \tau'}(n))$  is a solution of  $G$ . From this it follows that if  $G$  has a *unique* solution, then  $\bar{v}_{\text{opt}} = \bar{v}_{\sigma', \tau'}$ , for any pair of optimal strategies  $\sigma', \tau'$ .

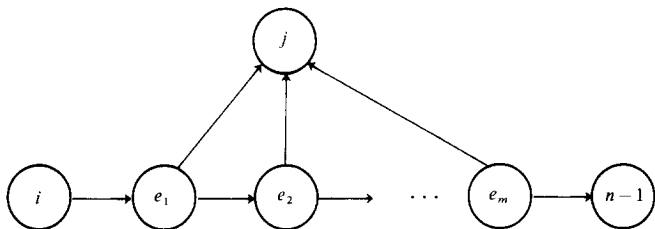
### 3. COMPLEXITY RESULTS

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

Therefore, we show that in a 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  $G' > \frac{1}{2}$  if and only if the value of  $G > \frac{1}{2}$ . The new game  $G'$  is called a *stopping SSG* because it halts with probability 1; this fact will be used to prove that  $G'$  has a unique solution. (The term “stopping stochastic game” was introduced by Shapley (1953). 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  $\{1, 2, \dots, n\}$  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  $\{e_1, e_2, \dots, e_m\}$ . 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  $G'$ :  $\{(i, e_1), (e_1, e_2), (e_2, e_3), \dots, (e_{m-1}, e_m), (e_m, n-1), (e_1, j), (e_2, j), \dots, (e_m, j)\}$ .



From this construction, note that if a path is followed from vertex  $i$ , the first vertex from the set  $\{1, \dots, n\}$  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  $\{1, \dots, n\}$  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  $\{1, \dots, n\}$ , the probability of ending the game before reaching another vertex in the set  $\{1, \dots, n-2\}$  is at least  $1/2^m$ .

Shapley (1953) showed that any stopping game constructed in this manner has a unique solution; we include the proof here for completeness.

**LEMMA 7** (Shapley, 1953). *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  $1, \dots, n$  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  $\{1, \dots, n\}$ , are uniquely determined by the values of the vertices in the set  $\{1, \dots, n\}$ . 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  $1 \leq i \leq n-2$ ,

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

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  $i$ th component of  $I(\bar{w})$  then

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

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  $> \frac{1}{2}$  if and only if the value of the corresponding  $(1/2^{cn})$ -stopping game is  $> \frac{1}{2}$ , for some constant  $c > 0$ . This lemma is a discrete version of a result of Filar (1981) who proved that the value of  $G$  is the limit, as  $\beta \rightarrow 0$ , of the value of the  $\beta$ -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  $> \frac{1}{2}$  if and only if the value of the corresponding  $(1/2^{cn})$ -stopping game is  $> \frac{1}{2}$ .*

*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  $G \leq \frac{1}{2}$  it must be the case that the value of any  $\beta$ -stopping game is  $\leq \frac{1}{2}$ .

Let  $G'$  be the  $\beta$ -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  $\sigma$  and  $\tau$  of the players, for any  $i \in \{1, \dots, n\}$ , the values of  $i$  with respect to  $\sigma$  and  $\tau$  in games  $G$  and  $G'$  are close. Specifically, for any vertex  $i \in \{1, \dots, n - 2\}$ ,  $v_{\sigma, \tau}(i) - v'_{\sigma, \tau}(i) \leq 1/4^n$  for sufficiently large  $c$ . Then from Lemma 2, if the value of  $G$  is greater than  $\frac{1}{2}$  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  $\{1, \dots, n-2\}$  have non-zero value in graphs  $G_{\sigma, \tau}$ ,  $G'_{\sigma, \tau}$ . Let these vertices be  $1, \dots, t$ ; let  $\bar{v}_{\sigma, \tau} = (v_{\sigma, \tau}(1), \dots, v_{\sigma, \tau}(t))$  and  $\bar{v}'_{\sigma, \tau} = (v'_{\sigma, \tau}(1), \dots, v'_{\sigma, \tau}(t))$ . Let  $\bar{v}_{\sigma, \tau} = Q\bar{v}_{\sigma, \tau} + \bar{b}$  be the equation uniquely defining the values  $v_{\sigma, \tau}(i)$ ,  $1 \leq i \leq t$ . 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

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

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

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  $j \rightarrow \infty$ . In the proof of Lemma 1, we showed that for all integers  $m \geq 0$ ,  $\|Q^m\| \leq (1 - 1/2^n)^m$ . From this it follows that

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

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  $\bar{v}_{\sigma, \tau}(\text{start}) > \frac{1}{2}$  then  $\bar{v}'_{\sigma, \tau}(\text{start}) > \frac{1}{2}$ .

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

**THEOREM 1.** *The SSG value problem is in  $NP \cap \text{co-}NP$ .*

*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  $> \frac{1}{2}$ .  $M$  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  $\bar{z}$  is a solution of  $G'$ . If so and if the value of the start vertex is  $> \frac{1}{2}$  then  $M$  accepts; else  $M$  rejects. From Lemma 8,  $M$  accepts  $G$  if and only if the value of  $G$  is  $> \frac{1}{2}$ .

A similar construction shows that the complement of the SSG value problem is in  $NP$ . A nondeterministic machine  $\bar{M}$ , given a SSG  $G$  as input, 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  $\bar{z}$  is a solution

of  $G'$ . If so and if the value of the start vertex is  $\leq \frac{1}{2}$  then  $\bar{M}$  accepts; else  $\bar{M}$  rejects. ■

To complete this section, we prove that special cases of the SSG value problem are actually in  $P$ .

**THEOREM 2.** *The SSG value problem restricted to SSGs with just (1) average and max vertices, or (2) average and min vertices, or (3) max and min vertices can be solved in polynomial time.*

*Proof.* The proof for case (1) is due to Derman (1972). Let  $G = (V, E)$  be an SSG with  $n$  vertices. From Lemma 8 we assume that  $G$  is a stopping game, so that  $G$  has a unique solution. We claim that the solution of  $G$  is the optimal solution to the following linear programming problem: minimize  $\sum_{l=1}^n v(l)$ , subject to the constraints

$$\begin{array}{ll} v(i) \geq v(j) & \text{if } i \text{ is a max vertex and } (i, j) \in E \\ v(i) \geq \frac{1}{2}(v(j) + v(k)) & \text{if } i \text{ is an average vertex and } (i, j), (i, k) \in E \\ v(i) = 0 & \text{if } i = n - 1 \\ v(i) = 1 & \text{if } i = n \\ v(i) \geq 0 & 1 \leq i \leq n. \end{array}$$

Let  $\bar{v} = (v(1), v(2), \dots, v(n))$  be the optimal solution to the linear programming problem. We first show that it must be the solution of the SSG  $G$ . Otherwise, one of two cases must arise. The first is that for some max vertex  $i$ ,  $(i, j)$  and  $(i, k)$  are edges of  $G$  but  $v(i) > v(k)$  and  $v(i) > v(j)$ . We construct a vector  $\bar{v}' = (v'(1), \dots, v'(n))$  such that  $\bar{v}'$  satisfies the constraints of the linear programming problem and  $\sum v'(i) < \sum v(i)$ , contradicting the fact that vector  $\bar{v}$  is an optimal solution to the linear programming problem. Let  $v'(l) = v(l)$  for  $l \neq i$  and let  $v'(l) = \max(v(j), v(k))$ . It is straightforward to verify that  $\bar{v}'$  satisfies the constraints of the linear programming problem and since  $v'(i) < v(i)$ ,  $\sum_{l=1}^n v'(l) < \sum_{l=1}^n v(l)$ , which proves the contradiction. The other case is that for some average vertex  $i$ ,  $(i, j)$  and  $(i, k)$  are edges of  $G$  but  $v(i) > \frac{1}{2}(v(j) + v(k))$ . A similar argument to the first case shows that this leads to a contradiction. Hence the optimal solution to the linear programming problem must be the unique solution of the SSG  $G$ .

Khachiyan (1979) has shown that the linear programming problem is computable in time polynomial in the length of the input, which is polynomial in  $n$  in this case, completing the proof for the max and average case.

The proof for case (2), where  $G$  has just min and average vertices, is very similar to the proof of (1). In this case, the linear programming problem is to maximize  $\sum_{l=1}^n v(l)$ , subject to the constraints

$v(i) \leq v(j)$	if $i$ is a min vertex and $(i, j) \in E$
$v(i) \leq \frac{1}{2}(v(j) + v(k))$	if $i$ is an average vertex and $(i, j), (i, k) \in E$
$v(i) = 0,$	if $i = n - 1$
$v(i) = 1,$	if $i = n.$

Finally, we consider case (3), where there are just max and min vertices. In this case, the following algorithm correctly computes the value of every node of  $G$  in polynomial time. The algorithm maintains two sets of vertices,  $U$  and  $D$ .  $D$  is the set of vertices whose values have already been determined by the algorithm and  $U$  is the set of vertices whose values are still undetermined. Initially  $D = \{n - 1, n\}$ , where the values  $v(n - 1)$  and  $v(n)$  are 0 and 1, respectively, since these are the sink vertices, and  $U = V - D$ . Throughout the algorithm,  $D \cup U = V$ . After the initialization step, the algorithm consists of the following loop:

```

repeat
  for all vertices  $i$  in  $U$  do
    if  $i$  is a max vertex with a 1-valued child in  $D$  then
      move  $i$  to  $D$  and define  $v(i)$  to be 1;
    if  $i$  is a max vertex with two 0-valued children in  $D$  then
      move  $i$  to  $D$  and define  $v(i)$  to be 0;
    if  $i$  is a min vertex with a 0-valued child in  $D$  then
      move  $i$  to  $D$  and define  $v(i)$  to be 0;
    if  $i$  is a min vertex with two 1-valued children in  $D$  then
      move  $i$  to  $D$  and define  $v(i)$  to be 1
  endfor
until no vertices are moved from  $U$  to  $D$  in the loop
for all vertices  $i$  in  $U$  do
  define  $v(i)$  to be 0.

```

The number of executions of the repeat loop is bounded by  $|V|$  since on every iteration of the repeat loop except the last, at least one vertex is moved from  $U$  to  $D$ . From this it is clear that the algorithm runs in polynomial time.

We next explain why the algorithm correctly computes the values of the vertices. Given a SSG  $G$  as input, let  $D_{\text{final}}$  and  $U_{\text{final}}$  be the sets  $U$  and  $D$  when the repeat loop of the algorithm is exited. It is straightforward to show that any vertex  $i$  in  $D_{\text{final}}$  has the correct value, by induction on the number of times the repeat loop is executed before  $i$  is moved to  $D$ .

Suppose vertex  $i$  is in  $U_{\text{final}}$ . To show that the algorithm is correct, we need to show that

$$\max_{\sigma} \min_{\tau} v_{\sigma, \tau}(i) = 0.$$



Let  $\tau$  be the following strategy. If  $i$  is a min vertex in  $D_{\text{final}}$  and both of its children have value 1, choose one child  $j$  arbitrarily and let  $\tau$  contain the edge  $(i, j)$ . Otherwise one of its children has value 0; choose a child  $j$  of  $i$  with value 0 and let  $\tau$  contain the edge  $(i, j)$ . If  $i$  is a min vertex in  $U_{\text{final}}$ , then one of its children, say  $j$ , is also in  $U_{\text{final}}$ . Otherwise  $i$  would have been moved from  $U$  to  $D$  at some iteration of the repeat loop. Again, let  $\tau$  contain the edge  $(i, j)$ . Now, let  $\sigma$  be any strategy of player 1. We show that  $v_{\sigma, \tau}(i) = 0$ .

To do this, it is sufficient to show that there is no path from  $i$  to the 1-sink in  $G_{\sigma, \tau}$ . Since there are no average vertices in  $G$ , the edges of  $G_{\sigma, \tau}$  are exactly the edges in  $\sigma \cup \tau$ . Suppose there is a path from  $i$  to the 1-sink. Then there must be a path from  $i$  to a vertex in  $D_{\text{final}}$  with value 1. This implies that there must be an edge of  $\sigma \cup \tau$  from some vertex in  $U_{\text{final}}$  to a vertex in  $D_{\text{final}}$  with value 1. But there is no edge from a max vertex of  $U_{\text{final}}$  to a vertex in  $D_{\text{final}}$  with value 1; otherwise  $i$  would have been moved from  $U$  to  $D$  on some iteration of the repeat loop. Also, there are no edges in  $\tau$  from a min vertex of  $U_{\text{final}}$  to a vertex of  $D_{\text{final}}$  with value 1, by definition of  $\tau$ . This contradicts the assumption that there is a path from  $i$  to the 1-sink in  $G_{\sigma, \tau}$ ; hence  $v(i) = 0$ . ■

### 3.1. Relationship between Stochastic Games and Complexity Classes

Our motivation for studying the SSG problem stems from a study of probabilistic complexity classes. The realization that a source of random bits can be useful in finding solutions to problems has motivated the study of probabilistic algorithms and models of computation. Recently, randomness has been added to nondeterministic and alternating machines in novel ways, giving rise to various “game-like” complexity classes. One example is the *games against nature* of Papadimitriou (1985), a model that combines the features of nondeterministic and probabilistic Turing machines (Gill, 1977). We briefly describe games against nature and some problems that they naturally model, and then show that the SSG problem with just max and average vertices is complete for the class of languages accepted by log space bounded games against nature. Finally, we show that the more general SSG value problem with max, min, and average vertices is complete for the class of languages accepted by log space bounded *randomized alternating Turing machines*.

Roughly, a game against nature is a special kind of nondeterministic Turing machine, whose states are partitioned into existential ( $\exists$ ) and random states. Without loss of generality, we assume that the Turing machine has exactly two choices at every step, except when a final state (accepting or rejecting) is reached, in which case the Turing machine halts. The nodes of a computation tree of such a machine on a fixed input are naturally

partitioned into existential and random nodes, depending on whether the state of the configuration labeling the node is existential or random, respectively.

Let  $M$  be a game against nature and let  $T$  be the computation tree of  $M$  on input  $x$ . Define an  $\exists$ -subtree  $T'$  to be any subtree of  $T$ , with the same root as  $T$ , where every internal existential node has exactly one child. The input  $x$  is accepted by  $M$  if and only if there is some  $\exists$ -subtree  $T'$  such that the probability of reaching an accepting leaf is  $> \frac{1}{2}$ , when a path is followed randomly from the root of  $T'$  by choosing each child of a random node with probability  $\frac{1}{2}$ . Time and space bounded games against nature can be defined just as are time and space bounded nondeterministic or probabilistic Turing machines.

Polynomial time bounded games against nature model a certain class of problems in optimization, called decision problems under uncertainty. An example of such a problem is the *stochastic scheduling problem*, described by Papadimitriou (1985). Suppose  $t$  tasks are to be scheduled on  $m$  processors to minimize the total time to execute all the tasks. There are precedence constraints among the tasks and the execution times of the tasks are random variables. The scheduling process can be thought of as a game against nature where player 1, modeling the scheduler, chooses the best set of tasks to schedule and player 0, modeling nature, randomly chooses which job will halt first. This problem is naturally modeled by a time bounded game against nature because the game between the scheduler and nature ends as soon as all jobs are completed. Papadimitriou used the model of a polynomial time bounded game against nature to show that the stochastic scheduling problem and similar decision problems under uncertainty are complete for PSPACE.

Other optimization problems, where long-term costs are to be optimized, can be modeled by *space bounded* games against nature. Derman (1972) describes an example of an inventory system for a given product where the inventory is under periodic review. After each review, a decision is made to add a certain amount of the product to the inventory level. The goal is to optimize a cost function that depends on factors such as the cost of holding the inventory in storage or the cost of shortages caused by random fluctuations in demand. A criterion for optimality would normally be a function of long-term costs.

Denote the class of log space bounded games against nature by GAN-SPACE( $\log(n)$ ). Condon (1989) showed that the SSG value problem, with just max and average vertices, is complete for the class GAN-SPACE( $\log(n)$ ). To see this, note that a game against nature can simulate a play of a SSG  $G$  using only log space, by following the edges of  $G$ , choosing an edge existentially from max vertices and randomly from average vertices. From this it follows that the SSG-value problem with max and average vertices is in GAN-SPACE( $\log(n)$ ). The problem is complete

for  $\text{GAN-SPACE}(\log(n))$ , with respect to deterministic log space reductions, because a space bounded game against nature  $M$  on a fixed input  $x$  can be represented by a SSG in the following way. The vertices of the SSG are labeled by configurations of  $M$  on  $x$  and the edges represent one-step transitions between configurations. If a vertex is labeled by a configuration with a random state, it is an average vertex, otherwise it is a max vertex. The vertices labeled by the initial, accept, and rejecting configurations are the start, the 1-sink, and the 0-sink vertices, respectively. The SSG constructed in this way has value  $> \frac{1}{2}$  if and only if the game against nature  $M$  accepts the fixed input  $x$ . Moreover, given a description of  $M$ , it is straightforward to construct the SSG for an input  $x$  in space  $O(\log |x|)$ .

Just as randomness can be added to nondeterministic machines to obtain games against nature, randomness can also be added to the alternating Turing machines of Chandra, Kozen, and Stockmeyer (1981) to obtain randomized alternating Turing machines. Randomized alternating Turing machines generalize games against nature in the following way: the states are partitioned into three types: existential ( $\exists$ ), universal ( $\forall$ ), and random. Let  $T$  be the computation tree of a randomized alternating Turing machine on input  $x$ . An  $\exists$ -subtree  $T'$  of  $T$  is defined just as before. We further define a  $\forall$ -subtree  $T''$  to be a subtree of  $T'$  with the same root as  $T'$ , where every internal universal node has exactly one child. The input  $x$  is accepted by  $M$  if there is some  $\exists$ -subtree  $T'$  such that for all  $\forall$ -subtrees  $T''$  of  $T'$ , the probability of reaching an accepting leaf of  $T''$  is  $> \frac{1}{2}$ , when a path is followed randomly from the root of  $T''$ . In (Condon, 1989), we denote the class of languages accepted by log space bounded randomized alternating Turing machines by  $\text{AUC-SPACE}(\log(n))$ . The SSG value problem is log space complete for the class  $\text{AUC-SPACE}(\log(n))$ . The proof of this is a straightforward generalization of the proof that the SSG problem with just max and average vertices is complete for  $\text{GAN-SPACE}(\log(n))$ .

#### 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  $0 \leq p \leq q \leq 2^m$ . 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  $NP \cap \text{co-}NP$ .

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 vertex with fan-out  $l$ , 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  $2' < q \leq 2'^{+1}$ , 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  $k$ ; and the remaining leaves point to  $i$ , the root of the tree. This construction guarantees that the probabilities of reaching  $j$  and  $k$  from  $i$  are 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 (1981). 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  $1 \leq p, q \leq 2^n$ . 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 are 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  $\sigma$  and  $\tau$  of the players, the value of the vertices of  $G$  with respect to  $\sigma$  and  $\tau$  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  $\sigma'$ ,  $\tau'$  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 SSGs with payoffs. Thus the results of Section 3 can be extended to show that the value problem for stopping SSGs with payoffs is in  $NP \cap \text{co-}NP$ .

The last generalization we consider was introduced by Shapley (1953) 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 binary values 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, 1953). However, it is an open problem whether the value problem for simultaneous stochastic games is in  $NP \cap \text{co-}NP$ . One obstacle to solving this is that the value of a simultaneous stochastic game need not be rational—see (Peters and Vrieze, 1987) 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 SSGs that have value  $> \frac{3}{4}$  and let  $B$  be the set of SSGs that have value  $< \frac{1}{4}$ . 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  $G \in A$ .

#### ACKNOWLEDGMENTS

I thank Laci 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. Thanks also to the careful comments of the anonymous referee, which greatly improved the presentation.

RECEIVED July 14, 1989; FINAL MANUSCRIPT RECEIVED June 27, 1990

#### REFERENCES

- CHANDRA, A. K., KOZEN, D. C., AND STOCKMEYER, L. J. (1981), Alternation, *J. Assoc. Comput. Mach.* **28** (1), 114–133.
- CONDON, A. (1989), “Computational Models of Games,” MIT Press, Cambridge, MA.
- DERMAN, C. (1972), “Finite State Markov Decision Processes,” Academic Press, New York.

- FILAR, J. A. (1981), Ordered field property for stochastic games when the player who controls transitions changes from state to state, *J. Algorithms* **34**, 503–515.
- GILL, J. (1977), The computational complexity of probabilistic Turing machines, *SIAM J. Comput.* **6**, 675–695.
- HOWARD, R. A. (1960), “Dynamic Programming and Markov Processes,” M.I.T. Press, Cambridge, MA.
- KEMENY, J. G., AND SNELL, J. LAURIE (1960), “Finite Markov Chains,” Van Nostrand, Princeton, N. J.
- KHACHIYAN, L. G. (1979), A polynomial algorithm in linear programming, *Soviet Math. Dokl.* **20**, 191–194.
- PAPADIMITRIOU, C. H. (1985), Games against nature, *J. Comput. System Sci.* **31**, 288–301.
- PETERS, H. J. M., AND VRIEZE, O. J. (1987), “Surveys in Game Theory and Related Topics,” CWI Tract 39, Centrum voor Wiskunde en Informatica, Amsterdam.
- SHAPLEY, L. S. (1953), Stochastic games, *Proc. Nat. Acad. Sci. U.S.A.* **39**, 1095–1100.