Cyclic Markov Equilibria in Stochastic Games[1]

JANOS FLESCH, FRANK THUIJSMAN AND KOOS VRIEZE

Department of Mathematics, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands

Abstract: We examine a three-person stochastic game where the only existing equilibria consist of cyclic Markov strategies. Unlike in two-person games of a similar type, stationary ε-equilibria ($\varepsilon > 0$) do not exist for this game. Besides we characterize the set of feasible equilibrium rewards.

1 Introduction

A two-person stochastic game can be described by a state space $S := \{1, \ldots, z\}$, and a corresponding collection $\{M_1, \ldots, M_z\}$ of bimatrices, where entry (i,j) of M_s consists of $r^1(s,i,j)$, $r^2(s,i,j) \in \mathbb{R}$ and a probability vector $(p(1|s,i,j), \ldots, p(z|s,i,j))$. The stochastic game is to be played in the following way. At each stage $n \in \mathbb{N}$ the play is in precisely one of the states. If the play is in state s at stage n then, simultaneously and independently, both players are to choose an action: player 1 chooses a row i of M_s, while player 2 chooses a column j of M_s. These choices induce an immediate payoff $r^1(s,i,j)$ to player 1 and $r^2(s,i,j)$ to player 2. Next, the play moves with probability $p(t|s,i,j)$ to state t, where new actions are to be chosen at stage $n + 1$. The description of n-person stochastic games, with $n \geq 3$, is analogous.

The players are assumed to have complete information and perfect recall. A player's strategy is a specification of a probability distribution, at each stage and state, over the available actions, conditional on the history of the play up to that stage. Strategies are generally denoted by π for player 1 and σ for player 2. A strategy is called stationary if, for each state, it specifies a mixed action to be used whenever this state is being visited. Stationary strategies are denoted by x and y. A stationary strategy is called pure, if for each state, it specifies one action to be chosen.

A pair of strategies (π, σ) with an initial state $s \in S$ determines a stochastic process on the payoffs. The sequences of payoffs are evaluated by the limiting

---

average reward, given for player $k \in \{1,2\}$ by

$$\gamma^k(s, \pi, \sigma) = E_{s\pi\sigma}\left(\liminf_{T \to \infty} \frac{1}{T} \sum_{n=1}^{T} R_n^k \right),$$

where R_n^k are random variables for the payoffs of player k at stage $n \in \mathbb{N}$.

A pair of strategies (π, σ) is a limiting average ε-equilibrium ($\varepsilon \geq 0$), if neither player can gain more than ε by unilateral deviation, i.e. if for all $s, \tilde{\pi}, \tilde{\sigma}$:

$$\gamma^1(s, \pi, \sigma) \geq \gamma^1(s, \tilde{\pi}, \sigma) - \varepsilon \quad \text{and} \quad \gamma^2(s, \pi, \sigma) \geq \gamma^2(s, \pi, \tilde{\sigma}) - \varepsilon.$$

A 0-equilibrium is simply called an equilibrium.

Generally in stochastic games, ε-equilibria do not always exist in terms of stationary strategies, and the question of existence among history dependent strategies is the most challenging open problem in stochastic game theory these days. This question has been affirmatively answered for many classes of two-person stochastic games. Due to special features in those classes, equilibria have often been achieved using techniques that do not apply for general n-person stochastic games. Many of these techniques are based on sequences of stationary equilibria in games, where either the strategy spaces are specifically restricted or the reward function is approximated by a continuous function (for example discounted rewards), that approach the original game in a certain sense. In many cases, one of the crucial properties of these sequences is that if one of the players, say player 1, plays his equilibrium strategy close to the limit and player 2 plays his limit strategy, then the reward of player 2 is individually rational, which enables player 1 to punish player 2's possible deviations.

In this paper we show that n-person stochastic games, with $n \geq 3$, require an analysis that is substantionally different from any analysis used for two-person games. This is done by examining a specific three-person stochastic game. To our knowledge, this is the first three-person stochastic game studied in detail. In this game, the gap between two-person and three-person stochastic games also appears in the nature of equilibria. This game is a recursive repeated game with absorbing states. A state is absorbing if any play that reaches this state will remain there permanently. A stochastic game is called recursive if all payoffs in the non-absorbing states are equal to zero. A repeated game with absorbing states is a stochastic game with only one non-absorbing state. Thus, a recursive repeated game with absorbing states is a repeated game with absorbing states where all the payoffs in the non-absorbing state are equal to zero. For two-person recursive repeated games with absorbing states Flesch et al. [1995] showed the existence of stationary ε-equilibria.
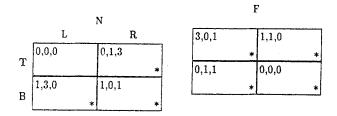
In the game presented below, no stationary ε-equilibria exist, so the two-person result does not extend to the n-person case. By the nature of this game, the players have only Markov strategies at their disposal. We show that the only equilibria in this game consist of Markov strategies that have a cyclic nature. As far as we know this is the first stochastic game where (cyclic) Markov strategies are

indispensable. So far Markov strategies mainly turned out to be important in finite horizon problems. This game also demonstrates that the class of stationary strategies, even by using threats in behavioral strategies, is too narrow to tackle the equilibrium existence problem for stochastic games with more than two players. Whether or not this class is rich enough for two-person stochastic games is not yet known. A thorough study of the potential possibilities of the use of (cyclic or non-cyclic) Markov strategies for infinite horizon stochastic games is needed.

In attempts to solve the existence problem of equilibria in stochastic games, a lot of attention has been devoted to special classes of stochastic games. For many of these classes the existence of stationary ε-equilibria could be derived (cf. Rogers [1969], Sobel [1971], Stern [1975], Filar [1984], Parthasarathy et al. [1984], Thuijsman & Vrieze [1991], Evangelista et al. [1994], Flesch et al. [1995]) or the existence of equilibria could be obtained using specifically selected stationary strategies and threats (cf. Vrieze & Thuijsman [1989], Thuijsman & Raghavan [1994], Thuijsman & Vrieze [1994]). These threats are often based on properties of solutions for two-person zero-sum stochastic games. Mertens & Neyman [1981] showed the existence of an undiscounted value by developing behavioral strategies where players use stationary strategies for history dependent periods of time.

To establish equilibria in two-person games one can often use that one player is securing a good reward for the other player while remaining above his own individually rational level. Such however is not possible in a three-person game since an increase of one players's reward may simultaneously be a decrease for another player, and with three players this may yield cyclic situations.

2 A Three-Person Example

We consider the game Γ:

In this cubic three-person game each player has two actions. The actions of the players are the following: player 1: Top, Bottom; player 2: Left, Right; player 3: Near, Far. The game is represented by taking separately the two layers of the cube

that belong to the two actions of player 3 (N and F). Absorbing entries are indicated by *'s, to be interpreted as: once such an entry is selected, the play moves to an absorbing state with these payoffs.

Note that all entries but one are absorbing, so the play absorbs as soon as one of the players chooses his second action, and also that the payoffs and the absorbing entries are cyclicly symmetric $(r^1(i_1, i_2, i_3) = r^2(i_3, i_1, i_2) = r^3(i_2, i_3, i_1)$ for each entry $(i_1, i_2, i_3))$. However we wish to emphasize that we have only introduced this cyclic symmetry to make the analysis of this game clearer. Existence of cyclic Markov equilibria and non-existence of stationary ε-equilibria, can also be obtained in non-symmetric games with the very same absorption structure.

In the game Γ, each probability distribution can be represented by the probability assigned to the second action, which lets the stationary strategy spaces equal $[0, 1]$ for each player. For stationary strategies of player 1, player 2 and player 3 we use the notations x, y and z respectively. The spaces of Markov strategies equal $[0, 1]^\infty$ for each player. Markov strategies are denoted by κ for player 1, by λ for player 2, and by μ for player 3.

For this game the only history up to stage $n \in \mathbb{N}$ if no absorption has occurred is the trivial one where each player has chosen his first action at all previous stages. Therefore all history dependent strategies are only Markov strategies. For a strategy triple $\theta = (\kappa, \lambda, \mu)$, let $q_\theta(i)$ denote the overall probability of absorption in entry $i = (i_1, i_2, i_3)$. Then the reward function is of the form

$$\gamma^k(\theta) = \sum_i q_\theta(i) r_i^k.$$

Notice that this expression holds irrespective whether $\sum_i q_\theta(i)$ is equal to or smaller than 1.

3   Analysis

Now we investigate the game Γ in detail. First we show that no stationary ε-equilibrium exists in Γ. For this purpose we shall use the well-known result that against fixed stationary strategies there always exists a pure stationary best reply (cf. Blackwell [1962], Hordijk et al. [1983]).

$$\forall y, z \exists x \in \{0, 1\} \forall \kappa [\gamma^1(x, y, z) \geq \gamma^1(\kappa, y, z)].$$

Obviously a similar statement is valid with respect to player 2 and player 3.

Lemma 3.1: There is no stationary equilibrium in Γ.

Proof: Suppose the opposite: let (x, y, z) be a stationary equilibrium. Recall that x, y, z are the probabilities on B, R and F respectively. First we prove that

$0 < x, y, z < 1$. If $x = 0$ then, because of a best reply argument, $y = 1$ and therefore $z = 0$, which contradicts $x = 0$. On the other hand $x = 1$ would imply $y = 0$, hence $z = 1$, which contradicts $x = 1$. So $0 < x < 1$, and by symmetry we also have $0 < y, z < 1$. Since $0 < x < 1$ we have

$$\frac{3(1-y)z + yz}{1 - (1-y)(1-z)} = \gamma^1(0, y, z) = \gamma^1(1, y, z) = 1 - z,$$

thus

$$y = \frac{z^2 + 2z}{z^2 + 1} > z.$$

By symmetry $z > x$ and $x > y$. Hence $y > z > x > y$, contradiction. □

Theorem 3.2: There is no stationary ε-equilibrium in Γ.

Proof: Suppose $(x^\varepsilon, y^\varepsilon, z^\varepsilon)$ is a stationary ε-equilibrium. We take a discrete sequence of $\varepsilon's$ converging to 0. Then by taking subsequences we can assume the sequence $(x^\varepsilon, y^\varepsilon, z^\varepsilon)$ to be convergent in the compact space $[0, 1]^3$. Let $(\tilde{x}, \tilde{y}, \tilde{z}) := \lim_{\varepsilon \downarrow 0} (x^\varepsilon, y^\varepsilon, z^\varepsilon)$. We distinguish two cases.

Case 1: $(\tilde{x}, \tilde{y}, \tilde{z})$ is absorbing, i.e. $\tilde{x} > 0$ or $\tilde{y} > 0$ or $\tilde{z} > 0$.
Suppose $\tilde{z} > 0$. Then $(x, \tilde{y}, \tilde{z})$ is also absorbing for all x. Because $(x^\varepsilon, y^\varepsilon, z^\varepsilon)$ is an ε-equilibrium we have

$$\gamma^1(x^\varepsilon, y^\varepsilon, z^\varepsilon) + \varepsilon \geq \gamma^1(x, y^\varepsilon, z^\varepsilon) \; \forall x.$$

Since

$$q_{(\tilde{x}, \tilde{y}, \tilde{z})}(i) = \lim_{\varepsilon \downarrow 0} q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(i), \qquad q_{(x, \tilde{y}, \tilde{z})}(i) = \lim_{\varepsilon \downarrow 0} q_{(x, y^\varepsilon, z^\varepsilon)}(i)$$

for each entry i, we obtain

$$\gamma^1(\tilde{x}, \tilde{y}, \tilde{z}) = \lim_{\varepsilon \downarrow 0} \gamma^1(x^\varepsilon, y^\varepsilon, z^\varepsilon) \geq \lim_{\varepsilon \downarrow 0} (\gamma^1(x, y^\varepsilon, z^\varepsilon) - \varepsilon) = \gamma^1(x, \tilde{y}, \tilde{z}) \forall x.$$

Similarly for player 2

$$\gamma^2(\tilde{x}, \tilde{y}, \tilde{z}) \geq \gamma^2(\tilde{x}, y, \tilde{z}) \forall y.$$

If $\tilde{x} > 0$ or $\tilde{y} > 0$ then similarly for player 3 as well

$$\gamma^3(\tilde{x}, \tilde{y}, \tilde{z}) \geq \gamma^3(\tilde{x}, \tilde{y}, z) \forall z,$$

otherwise

$$\gamma^3(\tilde{x}, \tilde{y}, \tilde{z}) = 1 \geq \gamma^1(\tilde{x}, \tilde{y}, z) \forall z.$$

Hence $(\tilde{x}, \tilde{y}, \tilde{z})$ is a stationary equilibrium, which contradicts Lemma 3.1.

Case 2: $(\tilde{x}, \tilde{y}, \tilde{z})$ is recurrent, i.e. $\tilde{x} = \tilde{y} = \tilde{z} = 0$.
By subsequence arguments and by symmetry we can assume without loss of generality that

$$q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(T, L, F) \geq \max \{ q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(T, R, N), q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(B, L, N) \} \forall \varepsilon.$$

We have

$$q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(T, L, F) = \frac{(1 - x^\varepsilon)(1 - y^\varepsilon) z^\varepsilon}{1 - (1 - x^\varepsilon)(1 - y^\varepsilon)(1 - z^\varepsilon)},$$

$$q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(B, L, N) = \frac{x^\varepsilon(1 - y^\varepsilon)(1 - z^\varepsilon)}{1 - (1 - x^\varepsilon)(1 - y^\varepsilon)(1 - z^\varepsilon)},$$

$$q_{(0, y^\varepsilon, z^\varepsilon)}(T, L, F) = \frac{(1 - y^\varepsilon) z^\varepsilon}{1 - (1 - y^\varepsilon)(1 - z^\varepsilon)}.$$

Then

$$\lim_{\varepsilon \downarrow 0} q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(B, L, N) > 0$$

would imply

$$\lim_{\varepsilon \downarrow 0} \gamma^1(x^\varepsilon, y^\varepsilon, z^\varepsilon) = \lim_{\varepsilon \downarrow 0} [3 q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(T, L, F) + q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(B, L, N)]$$

$$< \lim_{\varepsilon \downarrow 0} \left[3 \frac{q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(T, L, F)}{1 - q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(B, L, N)} \right]$$

$$= \lim_{\varepsilon \downarrow 0} [3(1 - x^\varepsilon) q_{(0, y^\varepsilon, z^\varepsilon)}(T, L, F)]$$

$$= \lim_{\varepsilon \downarrow 0} [3 q_{(0, y^\varepsilon, z^\varepsilon)}(T, L, F)]$$

$$= \lim_{\varepsilon \downarrow 0} \gamma^1(0, y^\varepsilon, z^\varepsilon),$$

so $(x^\varepsilon, y^\varepsilon, z^\varepsilon)$ would not be an ε-equilibrium for small ε. Hence,

$$\lim_{\varepsilon \downarrow 0} q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(B, L, N) = 0.$$

But then we get a contradiction:

$$\lim_{\varepsilon\downarrow 0} \gamma^2(x^\varepsilon, y^\varepsilon, z^\varepsilon) = \lim_{\varepsilon\downarrow 0} q_{(x^\varepsilon, y^\varepsilon, z^\varepsilon)}(T, R, N) < 1 = \lim_{\varepsilon\downarrow 0} \gamma^2(x^\varepsilon, 1, z^\varepsilon). \quad \Box$$

Now we turn to the class of Markov strategies. First we present a Markov equilibrium, which has a cyclic nature.

Theorem 3.3: Let Markov strategies κ, λ, μ for players 1,2,3 respectively be given by:

$$\kappa = (\tfrac{1}{2}, 0, 0, \tfrac{1}{2}, 0, 0, \tfrac{1}{2}, \ldots),$$

$$\lambda = (0, \tfrac{1}{2}, 0, 0, \tfrac{1}{2}, 0, 0, \ldots),$$

$$\mu = (0, 0, \tfrac{1}{2}, 0, 0, \tfrac{1}{2}, 0, \ldots).$$

Then (κ, λ, μ) is a Markov equilibrium in Γ with $\gamma(\kappa, \lambda, \mu) = (1, 2, 1)$.

Proof: First notice that

$$\gamma(\kappa, \lambda, \mu) = [\tfrac{1}{2}(1, 3, 0) + (\tfrac{1}{2})^2(0, 1, 3) + (\tfrac{1}{2})^3(3, 0, 1)] \cdot \sum_{t=0}^{\infty} (\tfrac{1}{2})^{3t} = (1, 2, 1).$$

We prove that κ is a best reply of player 1 against (λ, μ):

$$\gamma^1(\kappa, \lambda, \mu) \geq \gamma^1(\tilde{\kappa}, \lambda, \mu) \, \forall \tilde{\kappa}.$$

Similarly, neither player 2 nor player 3 has profitable deviations.

By way of contradiction suppose that a best reply $\tilde{\kappa} = (\tilde{x}_n)_{n=1}^{\infty}$ of player 1 against (λ, μ) is profitable, i.e. $\gamma^1(\tilde{\kappa}, \lambda, \mu) > \gamma^1(\kappa, \lambda, \mu)$ and $\gamma^1(\tilde{\kappa}, \lambda, \mu) \geq \gamma^1(\bar{\kappa}, \lambda, \mu) \, \forall \bar{\kappa}$ (a Markov best reply must exist because, due to the cyclic nature of λ and μ with periodicity three, the best replies are precisely the optimal solutions of a Markov decision problem with three non-absorbing states). We show that we need to have $\tilde{\kappa} = (0, 0, 0, \ldots)$. Let $\tilde{\kappa}_l := (\tilde{x}_n)_{n=l}^{\infty}$, and let $\kappa_l := (x_n)_{n=l}^{\infty}$ for all $l \in \mathbb{N}$. Let λ_l, μ_l be defined analogously. We have

$$\gamma^1(\tilde{\kappa}, \lambda, \mu) > \gamma^1(\kappa, \lambda, \mu) = 1.$$

So $\tilde{x}_1 = 0$, which implies

$$\gamma^1(\tilde{\kappa}_2, \lambda_2, \mu_2) = \gamma^1(\tilde{\kappa}, \lambda, \mu) > 1 = \gamma^1(\kappa_2, \lambda_2, \mu_2),$$

therefore $\tilde{x}_2 = 0$. Using the equality

$$\gamma^1(\tilde{\kappa}_2, \lambda_2, \mu_2) = \tfrac{1}{2}\gamma^1(\tilde{\kappa}_3, \lambda_3, \mu_3)$$

we obtain

$$\gamma^1(\tilde{\kappa}_3, \lambda_3, \mu_3) > 2 = \gamma^1(\kappa_3, \lambda_3, \mu_3),$$

so $\tilde{x}_3 = 0$. Now using

$$\gamma^1(\tilde{\kappa}_3, \lambda_3, \mu_3) = \tfrac{1}{2}3 + \tfrac{1}{2}\gamma^1(\tilde{\kappa}_4, \lambda_4, \mu_4)$$

we have

$$\gamma^1(\tilde{\kappa}_4, \lambda_4, \mu_4) > 1 = \gamma^1(\kappa_4, \lambda_4, \mu_4),$$

which yields $\tilde{x}_4 = 0$. This argument can be repeated this way, so indeed $\tilde{\kappa} = (0, 0, 0, \dots)$. But then we get a contradiction:

$$\gamma^1(\tilde{\kappa}, \lambda, \mu) = 1 = \gamma^1(\kappa, \lambda, \mu). \quad \square$$

Observe that for all $l \in \mathbb{N}$ the strategies $\kappa_l := (x_n)_{n=l}^\infty, \lambda_l := (y_n)_{n=l}^\infty, \mu_l := (z_n)_{n=l}^\infty$ form cyclic Markov equilibria as well. Also, if for $x \in [0, 1]$ and $n \in \mathbb{N}$ the notation $x(n)$ represents playing x for n subsequent stages, then

$$\pi_n = (\alpha(n), 0(n), 0(n), \alpha(n), 0(n), 0(n), \alpha(n), \dots)$$

$$\sigma_n = (0(n), \alpha(n), 0(n), 0(n), \alpha(n), 0(n), 0(n), \dots)$$

$$\tau_n = (0(n), 0(n), \alpha(n), 0(n), 0(n), \alpha(n), 0(n), \dots)$$

form an equilibrium for each n, where $(1 - \alpha)^n = \tfrac{1}{2}$. The equality $(1 - \alpha)^n = \tfrac{1}{2}$ makes that in any period of n stages the play absorbs with probability $\tfrac{1}{2}$.

The next Theorem says that all Markov equilibria are of the same type as presented in Theorem 3.3. Let (κ, λ, μ) be an equilibrium. Notice that we can assume that at any stage at least one player plays his second action with positive probability, otherwise this stage could be skipped without loosing that (κ, λ, μ) is an equilibrium. Formally therefore, we suppose $x_n + y_n + z_n > 0 \, \forall n \in \mathbb{N}$.

Theorem 3.4: Let (κ, λ, μ) be an equilibrium in Γ. Then, at each stage exactly one of the players plays his second action with positive probability, and these players appear cyclically in the order $1, 2, 3$.

Proof: In this proof we use the following notation:

$$(u_n, v_n, w_n) := \gamma(\kappa_n, \lambda_n, \mu_n) \, \forall n \in \mathbb{N}.$$

Step 1: $x_n, y_n, z_n < 1$ and $(\kappa_n, \lambda_n, \mu_n)$ is an equilibrium $\forall n \in \mathbb{N}$.
▷ If $x_1 = 1$ or $y_1 = 1$ or $z_1 = 1$, then (x_1, y_1, z_1) would be a stationary equilibrium, which would contradict Lemma 3.1. Hence stage 2 is reached with positive

probability, so $(\kappa_2, \lambda_2, \mu_2)$ is an equilibrium. Therefore $x_2, y_2, z_2 < 1$ again. This argument can be repated this way, so the proof is complete.

Step 2: $\exists n \in \mathbb{N}$ for which $x_n = 0$ or $y_n = 0$ or $z_n = 0$.
▷ Suppose the opposite: $0 < x_n, y_n, z_n < 1 \ \forall n \in \mathbb{N}$. Then for stage 1 we have

$$u_1 = \gamma^1((0, x_2, x_3, \ldots), \lambda, \mu) = \gamma^1((1, x_2, x_3, \ldots), \lambda, \mu),$$

hence

$$u_1 = 3(1 - y_1)z_1 + y_1 z_1 + (1 - y_1)(1 - z_1)u_2 = 1 - z_1.$$

By expressing u_2

$$u_2 = \frac{1 - 4z_1 + 2y_1 z_1}{(1 - y_1)(1 - z_1)}.$$

Similar equations hold concerning the other two players. Due to symmetry we can assume for stage 1 that $u_1 \leq \min\{v_1, w_1\}$. Then by the equations $u_1 = 1 - z_1$, $v_1 = 1 - x_1, w_1 = 1 - y_1$ we obtain $z_1 \geq \max\{x_1, y_1\}$. This implies

$$u_2 \leq 1 - \frac{z_1}{1 - z_1},$$

and then

$$u_1 - u_2 \geq \frac{z_1}{1 - z_1} - z_1 > z_1^2.$$

So we have

$$\min\{u_2, v_2, w_2\} \leq u_2 < u_1 - z_1^2 = \min\{u_1, v_1, w_1\} - (\max\{x_1, y_1, z_1\})^2$$
$$< \min\{u_1, v_1, w_1\}.$$

For stage 2 we have $u_2 = 1 - z_2, v_2 = 1 - x_2, w_2 = 1 - y_2$, which yields

$$\max\{x_2, y_2, z_2\} > \max\{x_1, y_1, z_1\}.$$

Then analogously

$$\min\{u_3, v_3, w_3\} < \min\{u_2, v_2, w_2\} - (\max\{x_2, y_2, z_2\})^2$$
$$< \min\{u_1, v_1, w_1\} - ((\max\{x_1, y_1, z_1\})^2 + (\max\{x_2, y_2, z_2\})^2),$$

and

$$\max\{x_3, y_3, z_3\} > \max\{x_2, y_2, z_2\}.$$

Using this inductively we find that as n increases the number $\min\{u_n, v_n, w_n\}$ goes below zero, which is a contradiction.

Step 3: If $z_n = 0$, then either $x_n = 0$ or $y_n = 0$.

▷ Assume by way of contradiction that $0 < x_n, y_n < 1$. Then we have

$$u_n = 1, u_{n+1} = \frac{u_n}{1 - y_n} > 1$$

$$v_n = 1 - x_n, v_{n+1} = \frac{v_n - 3x_n}{1 - x_n} < 1.$$

Since $u_{n+1} > 1$ we obtain $x_{n+1} = 0$. But then $v_{n+1} \geq 1$, contradiction.

Step 4: If $x_n > 0, y_n = z_n = 0$ then either $x_{n+1} > 0$, $y_{n+1} = z_{n+1} = 0$ or $y_{n+1} > 0$, $x_{n+1} = z_{n+1} = 0$.
▷ Since $1 \leq w_n = (1 - x_n)w_{n+1}$ we have $w_{n+1} > 1$. The second action of any player cannot give himself more than 1, so $z_{n+1} = 0$, and by step 3 either $x_{n+1} = 0$ or $y_{n+1} = 0$.

Step 5: If $x_n > 0$ and $y_n = z_n = 0$ then $\min\{u_n, v_n, w_n\} = 1$.
▷ Obviously, we have $u_n = 1$ and $w_n \geq 1$. Suppose $\bar{n}$ is the first stage after state n with $y_{\bar{n}} > 0$ (there must be such a stage, otherwise by step 4 we would obtain $z_n = z_{n+1} = \cdots = 0$, and hence $\gamma^3(\kappa_n, \lambda_n, \mu_n) = 0 < 1 = \gamma^3(\kappa_n, \lambda_n, 1)$ would hold contradicting the fact that (κ, λ, μ) is an equilibrium). Thus $\gamma^2(\kappa_{\bar{n}}, \lambda_{\bar{n}}, \mu_{\bar{n}}) = 1$. By step 4 we have $z_n = \cdots = z_{\bar{n}-1} = 0$. Therefore

$$v_n = 3(x_n + (1 - x_n)x_{n+1} + \cdots + (1 - x_n)\cdots(1 - x_{\bar{n}-2})x_{\bar{n}-1})$$

$$+ (1 - x_n)\cdots(1 - x_{\bar{n}-1})\gamma^2(\kappa_{\bar{n}}, \lambda_{\bar{n}}, \mu_{\bar{n}}),$$

and by step 1 we obtain $(1 - x_n)\cdots(1 - x_{\bar{n}-1}) > 0$, so $v_n > 1$.

Step 6: $x_1 = 0$ or $y_1 = 0$ or $z_1 = 0$.
▷ Suppose that n is the first stage when $x_n = 0$ or $y_n = 0$ or $z_n = 0$. If $n = 1$ then we are done. Otherwise assume by way of contradiction that $n > 1$. Then $0 < x_1, y_1, z_1, \ldots, x_{n-1}, y_{n-1}, z_{n-1} < 1$, and $u_1 = 1 - z_1 < 1$, and therefore just like above we have $\min\{u_n, v_n, w_n\} < \min\{u_1, v_1, w_1\} \leq u_1 < 1$, which contradicts the fact that $\min\{u_n, v_n, w_n\} = 1$ (cf. step 5).

Due to symmetry and step 6 we may suppose without loss of generality that $z_1 = 0$. Then by step 3 we have $x_1 = 0$ or $y_1 = 0$. Assume $y_1 = 0$, so $x_1 > 0$. By step 4 either $x_2 > 0, y_2 = z_2 = 0$ or $y_2 > 0, x_2 = z_2 = 0$. Now using step 4 inductively we obtain that at any stage exactly one of the players plays his second action with positive probability. And finally, the second statement of the theorem is an immediate consequence of step 4, so the proof is complete. □

Theorem 3.5: The set of feasible equilibrium rewards for Γ is

$$\Psi := \{(u, v, w) \in \mathbb{R}^3 \mid u, v, w \geq 1, u + v + w = 4, u = 1 \text{ or } v = 1 \text{ or } w = 1\}.$$

Proof: Let (κ, λ, μ) be a Markov equilibrium for Γ with rewards $(u, v, w) = \gamma(\kappa, \lambda, \mu)$. We show that $(u, v, w) \in \Psi$. Suppose $x_1 > 0$. Then by Theorem 3.4 we have $y_1, z_1 = 0$. Hence $u = 1, w \geq 1$. Let n be the first stage when $y_n > 0$. This implies that $\gamma^2(\kappa_n, \lambda_n, \mu_n) = 1$ and $z_1 = \cdots = z_{n-1} = 0$. Thus

$$v = 3(x_1 + (1 - x_1)x_2 + \cdots + (1 - x_1)\cdots(1 - x_{n-2})x_{n-1})$$
$$+ (1 - x_1)\cdots(1 - x_{n-1})\gamma^2(\kappa_n, \lambda_n, \mu_n),$$

and since $x_1, \ldots, x_{n-1} < 1$ (cf. proof of 3.4) we have $(1 - x_1)\cdots(1 - x_{n-1}) > 0$, so $v \geq 1$. The equality $u + v + w = 4$ is trivial.

Now we show that if $(u, v, w) \in \Psi$ then there exists a Markov equilibrium with these rewards. By symmetry it suffices to find a Markov equilibrium with rewards $(1, 1 + \alpha, 2 - \alpha)$, where $\alpha \in [0, 1]$. Let

$$\tilde{\kappa} = \left(\frac{\alpha}{2}, 0, 0, \frac{1}{2}, 0, 0, \frac{1}{2}, 0, \ldots\right)$$

(only the mixed action for stage 1 is α-dependent), and let λ and μ as defined in Theorem 3.3. Now $\gamma(\tilde{\kappa}, \lambda, \mu) = (1, 1 + \alpha, 2 - \alpha)$, and it can be verified similarly to the proof of Theorem 3.3 that $(\tilde{\kappa}, \lambda, \mu)$ is a Markov equilibrium indeed. $\square$

Finally, we wish to remark that a similar analysis can be succesfully employed to show the existence of ε-equilibria in any $2 \times 2 \times 2$ recursive repeated game with absorbing states which, like the example, has only one non-absorbing entry.

References

Blackwell D (1962) Discrete dynamic programming. Ann Math Statist 33: 719–726

Evangelista FS, Raghavan TES, Vrieze OJ (1994) ARAT repeated games with absorbing states. Mimeo, University of Illinois at Chicago

Filar JA (1984) On stationary equilibria of a single controller stochastic game. Mathematical Programming 30: 313–325

Flesch J, Thuijsman F, Vrieze OJ (1995) Recursive repeated games with absorbing states. Report M95-05, University of Limburg, Maastricht

Hordijk A, Vrieze OJ, Wanrooij GL (1983) Semi-Markov strategies in stochastic games. Int J Game Theory 12: 81–89

Mertens JF, Neyman A (1981) Stochastic games. Int J Game Theory 10: 53–66

Parthasarathy T, Tijs SH, Vrieze OJ (1984) Stochastic games with state independent transitions and separable rewards. In: Hammer G, Palaschke D (eds.) Selected Topics in Operations Research and Mathematical Economics, Springer Verlag, Berlin 262–271

Rogers PD (1969) Non-zerosum stochastic games. PhD thesis, report ORC 69-8, Operations Research Center, University of California, Berkeley

Sobel MJ (1971) Noncooperative stochastic games. Ann Math Statist 42: 1930–1935

Stern MA (1975) On stochastic games with limiting average payoff. PhD thesis, University of Illinois at Chicago

Thuijsman F, Raghavan TES (1994) Stochastic games with switching control or ARAT structure. Report M94-06, University of Limburg, Maastricht
Thuijsman F, Vrieze OJ (1991) Easy initial states in stochastic games. In: Raghavan TES et al (eds) Stochastic Games and Related Topics. Kluwer Academic Publishers, Dordrecht 85–100
Thuijsman F, Vrieze OJ (1994) The power of threats in stochastic games. Report M94-05, University of Limburg, Maastricht
Vrieze OJ, Thuijsman F (1989) On equilibria in repeated games with absorbing states. Int J Game Theory 18: 293–310