Randomized Strategies and Terminal Distributions

A. MAITRA W. SUDDERTH*

AMS 1991 Subject Classifications. 60G40, 93E20, 04A15, 28A12.
Key words and phrases.Measurable gambling, stochastic control, randomization, stopping time, Markov chain, the Skorohod stopping problem, analytic sets, saturated houses.

1 Introduction

A pure strategy in a Dubins and Savage (1976) gambling problem can be thought of as a sequence of gambles selected one-by-one from those available at each stage of the game. There are two natural ways to select a strategy at random. One method is to select the individual gambles at random at each stage. Another method is to choose the entire strategy at random from the set of all pure strategies. Our first result (Theorem 2.1) is that these two methods are equivalent in the context of measurable gambling theory. This result is related to similar results in game theory due to Kuhn (1953) and Aumann (1964), and in Markov decision processes due to Dynkin and Yushkevich (1979) and Feinberg (1996).

Our second theorem concerns the set of all possible terminal distributions that can be obtained by stopping a given Markov chain at random. Again we consider two ways to randomize - first by choosing a distribution at random from those terminal distributions that can be obtained using a nonrandom stopping time, and second by using a randomized stopping time. It turns out that the two methods lead to the same collection of terminal distributions (Theorem 4.4).

In the final section of the paper, we consider a third way to obtain a randomized terminal distribution. Namely, the decision to stop is made at random at each stage. Although it seems that the set of terminal distributions obtained should still be the same, our proof requires a further condition of some sort. Theorem 5.1 gives one such condition and another is explained in the discussion which follows its proof.

*Research supported by National Science Foundation Grant DMS-9703285.

2 Randomized strategies

Let X be a Borel subset of a Polish space and let $\mathbb{P}(X)$ be the set of probability measures on the Borel σ-field of X. Give $\mathbb{P}(X)$ the topology of weak convergence, so $\mathbb{P}(X)$ is again a Borel subset of a Polish space (see ([9], $17E$) for details). A *gambling house* on X is an analytic subset Γ of $X \times \mathbb{P}(X)$ such that each section $\Gamma(x)$ of Γ at x is nonempty. A *pure strategy σ available in Γ at x* is a sequence $\sigma_0, \sigma_1, \ldots$ such that $\sigma_0 \in \Gamma(x)$ and, for $n \geq 1, \sigma_n$ is a universally measurable function on X^n into $\mathbb{P}(X)$ such that $\sigma_n(x_1, x_2, \ldots, x_n) \in \Gamma(x_n)$ for every $x_1, x_2, \ldots, x_n \in X$. Such a pure strategy σ determines via the Ionescu-Tulcea Theorem ([3], Proposition 7.45) a unique probability measure on the Borel subsets of the history space $H = X^N$, where N is the set of positive integers and H is given the product topology. We will identify the strategy with the measure it induces and use the same symbol σ for the measure as well.

Let

$$\Sigma_\Gamma = \{(x, \sigma) \in X \times \mathbb{P}(H) \,:\, \sigma \,\text{is a pure strategy available in}\, \Gamma \,\text{at}\, x\}.$$

Because Γ is analytic, it follows that Γ has a universally measurable selector by virtue of the von Neumann selection theorem ([15], 29.9) and, consequently, each section $\Sigma_\Gamma(x)$ of Σ_Γ at x is nonempty. Furthermore, the set Σ_Γ is analytic as was proved by Dellacherie ([4], Theorem 3).

For a set $M \subseteq \mathbb{P}(X)$, denote by $\operatorname{sco} M$ the strong convex hull of the set M, that is, the set of $\mu \in \mathbb{P}(X)$ such that for some $\nu \in \mathbb{P}(\mathbb{P}(X))$ with $\nu^*(M) = 1$,

$$\mu(\cdot) = \int \lambda(\cdot)\, \nu(d\lambda).$$

A *behavior strategy* τ available in Γ at x is a pure strategy available in the gambling house $\operatorname{sco}\Gamma$ at x, where $(\operatorname{sco}\Gamma)(x) = \operatorname{sco}\Gamma(x)$. It is known that the operation "sco" preserves analytic sets ([5],p.186), so that $\operatorname{sco}\Gamma$ is a gambling house on X. As before identify τ with the probability measure it induces on the Borel subsets of H and use the same symbol τ for the measure. With this identification, the set of behavior strategies available in Γ at x is just $\Sigma_{\operatorname{sco}\Gamma}(x)$.

A *randomized strategy* ρ available in Γ at x is a probability measure on the Borel subsets of $\mathbb{P}(H)$ such that $\rho(\Sigma_\Gamma(x)) = 1$. Plainly, ρ induces in an obvious way a probability measure on the Borel subsets of H, namely, an element of $\operatorname{sco}\Sigma_\Gamma(x)$. We will therefore identify the set of randomized strategies available in Γ at x with $\operatorname{sco}\Sigma_\Gamma(x)$. For countable X, the house $\operatorname{sco}\Gamma$ was used by Hill and Pestien (1987) in their study of Markov strategies.

The main result of this section is that behavior strategies and randomized strategies induce the same set of distributions on H. More precisely, in the notation introduced above, we have:

Theorem 2.1. *For each $x \in X$, $\Sigma_{\text{sco}\,\Gamma}(x) = \text{sco}\,\Sigma_\Gamma(x)$.*

A simple consequence of this theorem is that a gambler gains nothing through randomization if his payoff is the integral of a bounded, Borel function on the history space H. The theorem can also be applied to the strategies available to one of the players in an n-person stochastic game when the strategies of the other players are fixed.

The equivalence of randomized and behavior strategies has been investigated in game theory by Kuhn (1953) and Aumann (1964). In their formulations there is a technical difficulty in defining a randomized strategy. As Aumann has pointed out in [2] (see also B.V.Rao (1971)), there is no "natural" Borel structure on the space of pure strategies on which a probability measure can be defined to induce randomization. To get around this difficulty, Aumann induces the randomization externally by mapping a standard space like [0, 1] into the space of pure strategies and then using Lebesgue measure for the randomization. Aumann then defines a behavior strategy as a special kind of randomized strategy, where, roughly speaking, the choices of actions at different instants of time are independent of each other. The main result of Aumann (1964) is then that any randomized strategy can be realized by a behavior strategy.

Contrast this with the situation in gambling where the space of pure strategies is identified with an analytic set of probability measures and this set admits a "natural" Borel structure, namely, its Borel σ-field. Pure strategies can consequently be randomized intrinsically without the intervention of an auxiliary space. It is also worthy of note that both inclusions in Theorem 2.1 are nontrivial, whereas in Aumann's formulation, a behavior strategy is definitionally a randomized strategy.

There are also several results related to Theorem 2.1 in the literature on Markov decision processes. Indeed, Feinberg (1996) proved that any behavior strategy can be realized by a randomized strategy, and the converse was shown by Dynkin and Yushkevich (1979, pages 91-92). These results are for a Borel measurable setting under the assumption that a Borel selector exists. Here we are able to avoid this assumption by using universally measurable strategies. Some other related results can be found in Feinberg (1982, 1982a, 1986, 1991), Gikhman and Skorohod (1979), and Krylov (1965). In particular, Feinberg (1991) proves an analogue of Theorem 2.1 for any class of policies satisfying a certain "strong non-repeating condition."

We now introduce some notation. For $\mu \in \mathbb{P}(H)$ and $n \geq 0$, μ^n will denote the μ-probability distribution of the first $n+1$ coordinates $x_1, x_2, \ldots, x_{n+1}$ and $\mu(x_1, x_2, \ldots, x_n)$ will denote a version of the μ-conditional distribution of x_{n+1} given $x_1, x_2, \ldots, x_n$ which is jointly measurable in $\mu, x_1, x_2, \ldots, x_n$. Similarly, $\mu[x_1, x_2, \ldots, x_n]$ will denote a version of the μ-conditional distribution of $x_{n+1}, x_{n+2}, \ldots,$ given $x_1, x_2, \ldots, x_n$, which is jointly measurable in $\mu, x_1, x_2, \dots, x_n$. Such versions exist by virtue of ([18], Lemma 2.2).

3 The proof of Theorem 2.1

The inclusion $\operatorname{sco}\Sigma_\Gamma(x) \subseteq \Sigma_{\operatorname{sco}\Gamma}(x)$ has been established in ([19], Theorem 4.14).

To prove the reverse inclusion, we begin with a well known result (see, for example, Aumann ([1])).

Lemma 3.1. *Suppose that Y and Z are Borel subsets of Polish spaces. Let $Q(y, E)$ be a universally measurable transition function (Markov kernel) on $Y \times \mathcal{B}(Z)$, where $\mathcal{B}(Z)$ is the Borel σ-field of Z. Then there is a universally measurable function $\phi : Y \times [0,1] \to Z$ such that $Q(y, \cdot) = \lambda\phi_y^{-1}$ for all $y \in Y$, where λ is Lebesgue measure on $[0,1]$ and ϕ_y is the section of ϕ at y.*

The proof, which is routine, proceeds by identifying Z with $[0,1]$ and then uniformly "inverting" the distribution functions corresponding to the probability measures $Q(y,\cdot), y \in Y$.

We return to the proof of the reverse inclusion. Fix $x_0 \in X$ and let $\mu \in \Sigma_{\operatorname{sco}\Gamma}(x_0)$. By using the von Neumann selection theorem, one can define, for each $n \geq 0$, a universally measurable function $\tau_n : X^n \to \mathbb{P}(\mathbb{P}(X))$ such that

(a) $\tau_0(\Gamma(x_0)) = 1$,

(b) $\mu^0(\cdot) = \int \gamma(\cdot)\, \tau_0(d\gamma)$,

(c) $\tau_n(x_1, x_2, \dots, x_n)(\Gamma(x_n)) = 1$,
and $\mu(x_1, x_2, \dots, x_n)(\cdot) \quad = \int \gamma(\cdot)\, \tau_n(x_1, x_2, \dots, x_n)(d\gamma)$ for almost all (μ^{n-1}) $(x_1, x_2, \dots, x_n) \in X^n$.

By Lemma 3.1, choose, for each $n \geq 0$, a universally measurable function

$$\phi_n : X^n \times [0,1] \to \mathbb{P}(X)$$

such that $\tau_n(x_1, x_2, \dots, x_n) = \lambda\phi_n(x_1, x_2, \dots, x_n, \cdot)^{-1}$ for all $(x_1, x_2, \dots, x_n) \in X^n$. Set $H' = [0,1]^\omega$ and give H' the usual Borel σ-field and product Lebesgue measure λ^∞. With each $\vec{s} = (s_0, s_1, \dots, s_n, \dots) \in H'$, we associate a probability measure $\nu = \psi(\vec{s})$ on the Borel subsets of H as follows: the ν-distribution of x_1 is $\phi_0(s_0)$, and for $n \geq 1$ and $(x_1, x_2, \dots, x_n) \in X^n$, the ν-conditional distribution of x_{n+1} is $\phi_n(x_1, x_2, \dots, x_n, s_n)$. Then it is easily verified that ψ is universally measurable from H' to $\mathbb{P}(H)$ (see [3], p.177 for details). Let

$$m(A) = \int \psi(\vec{s})(A)\, \lambda^\infty(d\vec{s}) \tag{3.1}$$

for Borel subsets A of H.

We claim that $\mu = m$. To prove the claim, it will suffice to prove that $\mu^n = m^n$ for all $n \geq 0$. For $n = 0$ this follows from (3.1), the change of variable formula, the fact that $\tau_0 = \lambda\phi_0^{-1}$ and (b). For the inductive step, assume that $\mu^{n-1} = m^{n-1}$. Let A be a Borel subset of X^n and B a Borel subset of X. Now calculate as follows using the abbreviation $\vec{x}$ for $(x_1, x_2, \ldots, x_n)$:

$$
\begin{aligned}
& m^n(A \times B) \\
=& \int \cdots \int \int [\int_A \phi_n(\vec{x}, s_n)(B)\, \psi(s_0, s_1, \ldots, s_{n-1}, \ldots)^{n-1}\,(d\vec{x})]\, ds_n ds_{n-1} \ldots ds_0 \\
=& \int \cdots \int \{\int_A [\int \phi_n(\vec{x}, s_n)(B)\, ds_n] \psi(s_0, s_1, \ldots, s_{n-1}, \ldots)^{n-1}(d\vec{x})\} ds_{n-1} \ldots ds_0 \\
=& \int \cdots \int \{\int_A [\int \gamma(B)\, \tau_n(\vec{x})(d\gamma)] \psi(s_0, s_1, \ldots, s_{n-1}, \ldots)^{n-1}(d\vec{x})\} ds_{n-1} \ldots ds_0 \\
=& \int_A [\int \gamma(B)\, \tau_n(\vec{x})(d\gamma)]\, m^{n-1}(d\vec{x}) \\
=& \int_A [\int \gamma(B)\, \tau_n(\vec{x})(d\gamma)]\, \mu^{n-1}(d\vec{x}) \\
=& \int_A \mu(\vec{x})(B)\, \mu^{n-1}(d\vec{x}) \\
=& \mu^n(A \times B),
\end{aligned}
$$

where the third equality is a consequence of the change of variable theorem and the fact that $\tau_n(\vec{x}) = \lambda\phi_n(\vec{x}, \cdot)^{-1}$, the fifth is by virtue of the inductive hypothesis, and the sixth is by (c). It follows that $\mu^n = m^n$.

We set $\rho = \lambda^\infty \psi^{-1}$. Since $\mu = m$, it will follow from (3.1) and the change of variable theorem that $\mu \in \operatorname{sco} \Sigma_\Gamma(x_0)$ as soon as we show that $\rho(\Sigma_\Gamma(x_0)) = 1$. Towards establishing this, observe that

$$\psi(\vec{s}) \in \Sigma_\Gamma(x_0)$$

if and only if

$$\psi(\vec{s})^0 \in \Gamma(x_0)$$

and ,for all $n \geq 1$,

$$\psi(\vec{s})^{n-1}(\{(x_1, x_2, \ldots, x_n) \in X^n \,:\, \psi(\vec{s})(x_1, x_2, \ldots, x_n) \in \Gamma(x_n)\}) = 1.$$

It will therefore suffice to show that each of the sets following the "if and only if" has λ^∞-measure one. First

$$\lambda^\infty(\{\vec{s} \in H' \,:\, \psi(\vec{s})^0 \in \Gamma(x_0)\}) = \lambda(\{s_0 \in [0,1] \,:\, \phi_0(s_0) \in \Gamma(x_0)\} = 1$$

because $\lambda\phi_0^{-1} = \tau_0$ and $\tau_0(\Gamma(x_0)) = 1$. Next note that, for each n, the set of $\vec{s}$ satisfying the last condition above can be written as

$$E=\{\vec{s}\in H' : \psi(\vec{s})^{n-1}(\{(x_1, x_2, \dots, x_n)\in X^n : \phi_n(x_1, x_2, \dots, x_n, s_n)\in\Gamma(x_n)\})=1\}.$$

It is now evident that the set E is determined by the first $n+1$ coordinates of $\vec{s}$. Fix $s_0, s_1, \dots, s_{n-1}$. In order to prove that $\lambda^\infty(E) = 1$, it will suffice to show that the $(s_0, s_1, \dots, s_{n-1})$-section of E has Lebesgue measure one. Abbreviate the measure $\psi(s_0, s_1, \dots, s_{n-1}, \dots)^{n-1}$ by η and consider the set

$$D = \{(x_1, x_2, \dots, x_n, s_n) \in X^n \times [0,1] : \phi_n(x_1, x_2, \dots, x_n, s_n) \in \Gamma(x_n)\}.$$

Recall that $\lambda\phi_n(x_1, x_2, \dots, x_n, \cdot)^{-1} = \tau_n(x_1, x_2, \dots, x_n)$ and $\tau_n(x_1, x_2, \dots, x_n)$ $(\Gamma(x_n)) = 1$. It follows that each $(x_1, x_2, \dots, x_n)$-section of D has Lebesgue measure one, hence $(\eta \times \lambda)(D) = 1$. Consequently,

$$\lambda(\{s_n \in [0,1] : \eta(D_{x_1, x_2, \dots, x_n}) = 1\}) = 1.$$

But the above set is just the $(s_0, s_1, \dots, s_{n-1})$-section of E. So the proof of Theorem 2.1 is complete.

4 Terminal gambles

Cosider a gambling house Γ on a Borel subset X of a Polish space such that $\Gamma(x)$ is a singleton set $\{\gamma_x\}$ for each $x \in X$. It follows that the map $x \to \gamma_x$ is Borel measurable. For each $x \in X$, there is a unique strategy $\sigma(x)$ available in Γ at x. Under $\sigma(x)$, the coordinate random variables $X_n, n \geq 1$, on H form a Markov chain with initial state $X_0 = x$ and stationary transition probability function $x \to \gamma_x$.

Terminal gambles are distributions on the state space X obtained by stopping the Markov chain. To make this precise, recall that a *stopping time* is a universally measurable function on H into $\bar{\omega} = \{0, 1, 2, \dots\} \cup \{\infty\}$ such that whenever $t(h) = n < \infty$ and h' agrees with h through the first n coordinates, $t(h) = n$. If t is a stopping time such that $\sigma(x)(\{t < \infty\}) = 1$, denote by $\gamma(t)$ the $\sigma(x)$-distribution of X_t and let

$$\begin{aligned}\Gamma^{pc} =\{&(x, \gamma) \in X \times \mathbb{P}(X) : \gamma = \gamma(t) \text{ for some stopping time } t \\ &\text{such that } \sigma(x)(\{t < \infty\}) = 1\}.\end{aligned}$$

It turns out Γ^{pc} is itself a gambling house (see Lemma 4.1) and is called the *pseudo-composition closure* of Γ. (The term "pseudo" is used here because Dubins and Savage (1976) used "composition closure" to denote a gambling house defined as above except that they required that their stopping times be everywhere finite.) For each $x \in X$, $\Gamma^{pc}(x)$ is a set of terminal gambles obtained by stopping the Markov chain, starting at x, at different stopping times. Plainly, a larger set of terminal gambles will arise if we allow randomized stopping times. A *randomized stopping time* τ is a universally measurable function on $[0,1] \times H$ into $\bar{\omega}$ such that for each $z \in [0,1]$, the section τ_z is a stopping time on H. As before, if τ is a randomized stopping time such that $\lambda \times \sigma(x)(\{\tau < \infty\}) = 1$, let $\gamma(\tau)$ be the $\lambda \times \sigma(x)$-distribution of X_τ and let

$$\Gamma^{ran} = \{(x,\gamma) \in X \times \mathbb{P}(X) \,:\, \gamma = \gamma(\tau) \text{ for some randomized stopping time } \tau \text{ such that } \lambda \times \sigma(x)(\{\tau < \infty\}) = 1\},$$

where λ is Lebesgue measure.

The main result of this section (Theorem 4.4) is that for each $x \in X$, $\Gamma^{ran}(x)$ is the strong convex hull of $\Gamma^{pc}(x)$. In other words, for a Markov chain, we obtain exactly the same terminal gambles by randomization of stopping times at the outset as by randomizing after obtaining the terminal gambles from (deterministic) stopping times.

The gambling house Γ^{ran} (see Lemma 4.1 and Theorem 4.4) can also be described, courtesy of a result of H.Rost ([5], p.50), in the following manner. Recall that a bounded, Borel measurable function g on X is Γ'-*excessive*, where Γ' is a gambling house on X, if $\int g\, d\gamma \leq g(x)$ for every $\gamma \in \Gamma'(x)$ and $x \in X$. The result of Rost can be paraphrased thus: if Γ is a gambling house such that each $\Gamma(x)$ is a singleton, then Γ^{ran} is the largest gambling house on X with exactly the same excessive functions as Γ, that is, Γ^{ran} is the *saturation* of Γ. (The saturation is treated by Dellacherie and Meyer (1983) and again in [19].)

Lemma 4.1. *The set Γ^{pc} is analytic in $X \times \mathbb{P}(X)$. Consequently, Γ^{pc} is a gambling house on X.*

Proof. For each $x \in X$ let $\Delta(x)$ denote the probability measure on H that assigns mass 1 to the history $(x, x, \dots)$. Call a measure $\mu \in \mathbb{P}(H)$ *almost surely stagnant* if μ is the image of ν under the map $h \to (h_1, h_2, \dots, h_{t(h)}, h_{t(h)}, \dots)$ for some $\nu \in \mathbb{P}(H)$ and stopping time t such that $\nu(\{t < \infty\}) = 1$. Let $\mathbb{P}_{as}$ be the set of almost surely stagnant measures on H. By ([18], Lemma 9.4), $\mu \in \mathbb{P}_{as}$ if and only if

$$\mu(\{h \,:\, \mu[h_1, h_2, \dots, h_n] = \Delta(h_n)\}) \to 1 \text{ as } n \to \infty.$$

The condition above is clearly Borel in μ, so $\mathbb{P}_{as}$ is Borel in $\mathbb{P}(H)$, as was noted in [18].

For $\mu \in \mathbb{P}_{as}$ and $x \in X$, define

$$t^x_\mu = \inf\{n \geq 0 \,:\, \mu[h_1, h_2, \dots, h_n] = \Delta(h_n)\},$$

where, for $n = 0, h_n = x$, and $\mu[h_1, h_2, \dots, h_n] = \mu$. Then t^x_μ is a stopping time with $\mu(\{t^x_\mu < \infty\} = 1$. Let Σ_{as} be the subset of $X \times \mathbb{P}_{as}$ defined by the condition below:

$$\mu(\{h : \forall n(\text{either}\, \mu[h_1, h_2, \dots, h_n] = \Delta(h_n) \text{ or } \mu(h_1, h_2, \dots, h_n) = \gamma_{h_n})\}) = 1. \tag{4.1}$$

Plainly, Σ_{as} is a Borel subset of $X \times \mathbb{P}(H)$.

To complete the proof, let

$$f(h) = \begin{cases} \lim h_n & \text{, if the limit exists,} \\ x* & \text{, otherwise,} \end{cases} \tag{4.2}$$

where $x*$ is a fixed element of X. Then it is easy to see that

$$\Gamma^{pc} = \{(x, \mu f^{-1}) \in X \times \mathbb{P}(X) : (x, \mu) \in \Sigma_{as}\}. \tag{4.3}$$

So Γ^{pc} is a Borel measurable image of the Borel set Σ_{as}. Consequently, Γ^{pc} is analytic. □

Lemma 4.2. $\operatorname{sco}\Gamma^{pc} \subseteq \Gamma^{ran}$.

Proof. Let $\bar{\gamma} \in \operatorname{sco}(\Gamma^{pc}(x))$. So there is a probability measure $\nu \in \mathbb{P}(\mathbb{P}(X))$ such that $\nu(\Gamma^{pc}(x)) = 1$ and $\bar{\gamma}(B) = \int \gamma'(B)\, \nu(d\gamma')$ for each Borel subset B of X. Use (4.3) to fix a universally measurable function $g : \Gamma^{pc} \to \mathbb{P}(H)$ such that $(x, g(x, \gamma)) \in \Sigma_{as}$ and $g(x, \gamma)f^{-1} = \gamma$, where Σ_{as} is defined by the condition (4.1) and f by (4.2). Let ϕ be a Borel measurable function on $[0, 1]$ into $\mathbb{P}(X)$ such that $\lambda\phi^{-1} = \nu$, where λ is Lebesgue measure on $[0, 1]$.

Define τ on $[0, 1] \times H$ as follows:

$$\tau(z, h) = \begin{cases} \inf\{n \geq 0 : g(x, \phi(z))[h_1, h_2, \dots, h_n] = \Delta(h_n)\}, & \text{if}\, \phi(z) \in \Gamma^{pc}(x), \\ 0, & \text{otherwise.} \end{cases}$$

Then τ is a randomized stopping time.

Let P be the probability measure $\lambda \times \sigma(x)$. Then, for Borel $B \subseteq X$,

$$P(\{\tau < \infty, X_\tau \in B\} \mid Z = z) = \phi(z)(B) \text{ a.s.,}$$

so

$$\begin{aligned} P(\{\tau < \infty, X_\tau \in B\}) &= \int \phi(z)(B)\, \lambda(dz) \\ &= \int \gamma'(B)\, \lambda\phi^{-1}(d\gamma') \\ &= \int \gamma'(B)\, \nu(d\gamma') \\ &= \bar{\gamma}(B), \end{aligned}$$

where the second equality is by virtue of the change of variable theorem. It follows that $\bar{\gamma} \in \Gamma^{ran}(x)$. This completes the proof. □

Lemma 4.3. $\Gamma^{ran} \subseteq \operatorname{sco}\Gamma^{pc}$.

Proof. Suppose that $\gamma \in \Gamma^{ran}(x)$. So there is a randomized stopping time τ such that

$$\gamma(B) = P(\{\tau < \infty, X_\tau \in B\})$$

for every Borel subset B of X, where $P = \lambda \times \sigma(x)$. For $z \in [0,1]$, let γ_z be the terminal gamble under $\sigma(x)$ when the stopping time τ_z is used. Note that , for almost all z, $\gamma_z \in \Gamma^{pc}(x)$ and that the map $\psi(z) = \gamma_z$ is defined for almost all z and is universally measurable. Since, for any Borel subset B of X,

$$\begin{aligned}\gamma(B) &= \int \gamma_z(B)\,\lambda(dz) \\ &= \int \gamma'(B)\,\lambda\psi^{-1}(d\gamma')\end{aligned}$$

and $(\lambda\psi^{-1})(\Gamma^{pc}(x)) = 1$, it follows that $\gamma \in \operatorname{sco}(\Gamma^{pc}(x))$, which completes the proof. □

Putting Lemmas 4.2 and 4.3 together, we have

Theorem 4.4. $\Gamma^{ran} = \operatorname{sco}\Gamma^{pc}$.

5 Terminal gambles and behavior strategies

Let Γ be as in the previous section. We will continue with our study of the gambling house Γ^{ran}. Let Γ^l be the leavable closure of Γ, that is, $\Gamma^l(x) = \{\gamma_x, \delta(x)\}$, $x \in X$. (Here $\delta(x)$ is the probability measure that assigns mass one to $\{x\}$.) Plainly, Γ^l has the same excessive functions as Γ. So Γ^{ran} is the saturation of Γ^l as well as that of Γ. The question arises if every gamble in Γ^{ran} can be realized as a terminal gamble by using a deterministic stopping time and a behavior strategy in the house Γ^l, that is, the randomization is performed on the gambles in Γ^l, but not on the stopping times. We do not know the answer in general. In a special case, the answer is yes, and that is our next result.

Theorem 5.1. *If $\gamma_x(\{x\}) = 0$, for every $x \in X$, then*

$$\Gamma^{ran} \subseteq (\operatorname{sco}\Gamma^l)^{pc}.$$

Proof. Fix $x_0 \in X$ and let $\gamma \in \Gamma^{ran}(x_0)$. So there is a randomized stopping time τ such that $\lambda \times \sigma(x_0)(\{\tau < \infty\}) = 1$ and γ is the $\lambda \times \sigma(x_0)$-distribution of X_τ. Write P for $\lambda \times \sigma(x_0)$ and $Q(x, \cdot)$ for the transition function $x \to \gamma_x$. If μ is a finite measure on X, we define the measure μQ by

$$\mu Q(B) = \int Q(x, B)\, \mu(dx).$$

For each $n \geq 0$ and B a Borel subset of X, define

$$\gamma_n(B) = P(\{X_n \in B, \tau = n\})$$

and

$$\gamma_n'(B) = P(\{X_n \in B, \tau > n\}).$$

It is then easy to verify that

$$\gamma_0 + \gamma_0' = \delta(x_0) \tag{5.1}$$

and, for $n \geq 0$,

$$\gamma_n + \gamma_n' = \gamma_{n-1}' Q. \tag{5.2}$$

For each $n \geq 0$, fix a Radon-Nikodym derivative α_n of γ_n with respect to $\gamma_n + \gamma_n'$ such that $0 \leq \alpha_n \leq 1$.

Define now a behavior strategy σ^* available in Γ^l at x_0, or equivalently, a pure strategy available in $\operatorname{sco} \Gamma^l$ at x_0 as follows:

$$\sigma_0^* = \alpha_0(x_0)\delta(x_0) + (1 - \alpha_0(x_0))Q(x_0, \cdot)$$

and, for $n \geq 1$,

$$\sigma_n^*(x_1, x_2, \ldots, x_n) = \alpha_n(x_n)\delta(x_n) + (1 - \alpha_n(x_n))Q(x_n, \cdot).$$

We are going to run the process $X_n, n \geq 0$, with $X_0 = x_0$ according to σ^* and stop it at t^*, where the stopping time t^* is defined as follows:

$$t^*(h) = \begin{cases} \text{least } n \geq 1 \text{ such that } h_n = h_{n-1}, & \text{if such an n exists,} \\ \infty, & \text{otherwise.} \end{cases}$$

We now claim that (a) $\sigma^*(\{t^* < \infty\}) = 1$ and (b) the σ^*-distribution of X_{t^*} is γ. It will then follow that $\gamma \in (\operatorname{sco} \Gamma^l)^{pc}(x_0)$. Both claims follow immediately from

$$\sigma^*(\{t^* = n + 1, X_{n+1} \in B\}) = \gamma_n(B) \tag{5.3}$$

for every Borel subset B of X and $n \geq 0$.

A simple calculation shows that

$$\sigma^*(\{t^* = 1, X_1 \in B\}) = \alpha(x_0)1_B(x_0) = \gamma_0(B).$$

Let $n \geq 1$. It is easy to check that

$$\begin{aligned}&\sigma^*(\{t^* = n+1, X_{n+1} \in B\}) \\ &= (1 - \alpha_0(x_0))E_{\sigma(x_0)}(\prod_{i=1}^{n-1}(1 - \alpha_i(X_i))\alpha_n(X_n)1_B(X_n)).\end{aligned} \tag{5.4}$$

We use here the hypothesis that $\gamma_x(\{x\}) = 0$ for each $x \in X$.

On the other hand,

$$\begin{aligned}\gamma_n(B) &= \int_B \alpha_n(x_n)\,(\gamma_n + \gamma_n')(dx_n) \\ &= \int_B \alpha_n(x_n)(\gamma_{n-1}'Q)(dx_n) \quad \text{(from (5.2))} \\ &= \int\int_B \alpha_n(x_n)Q(x_{n-1}, dx_n)\gamma_{n-1}'(dx_{n-1}) \\ &= \int\int_B \alpha_n(x_n)(1 - \alpha_{n-1}(x_{n-1}))Q(x_{n-1}, dx_n)(\gamma_{n-1} + \gamma_{n-1}')(dx_{n-1}).\end{aligned}$$

If we iterate this process, using (5.2) repeatedly, we will obtain

$$\begin{aligned}\gamma_n(B) &= (1 - \alpha_0(x_0))E_{\sigma(x_0)}(\prod_{i=1}^{n-1}(1 - \alpha_i(X_i))\alpha_n(X_n)1_B(X_n)) \\ &= \sigma^*(\{t^* = n+1, X_{n+1} \in B\}) \qquad \text{(from (5.4))}\end{aligned}$$

So (5.3) is established and the proof is complete.

□

The inclusion $(\text{sco}\,\Gamma^l)^{pc} \subseteq \Gamma^{ran}$ is always true. This can be seen by using Rost's theorem, cited in the previous section, and the fact that the excessive functions for the gambling house $(\text{sco}\,\Gamma^l)^{pc}$ are exactly the same as the Γ-excessive functions.

Here is a way of eliminating the hypothesis from Theorem 5.1. But there is a price to pay in that we need to enlarge the state space. The details are as follows.

Let X' be a copy of X that is disjoint from X. If $x \in X$, x' will denote the copy of x in X'. Enlarge the state space X to $X \cup X'$ and define a new gambling house $\bar{\Gamma}$ on $X \cup X'$ as follows:

$$\bar{\Gamma}(x) = \{\gamma_x, \delta(x')\} \qquad \text{if } x \in X,$$

and

$$\bar{\Gamma}(x') = \{\delta(x)\} \qquad \text{if } x' \in X'.$$

Redefine σ^* and t^* as follows:

$$\sigma_0^* = \alpha_0(x_0)\delta(x_0') + (1 - \alpha_0(x_0))Q(x_0, \cdot)$$

and, for $n \geq 1$,

$$\sigma_n^*(x_1, x_2, \ldots, x_n) = \begin{cases} \alpha_n(x_n)\delta(x_n') + (1 - \alpha_n(x_n))Q(x_n, \cdot), & \text{if } x_n \in X \\ \delta(x), \quad \text{if } x_n = x' \in X'; \end{cases}$$

$$t^*(h) = \begin{cases} \text{least } n \geq 1 \text{ such that} h_{n-1} \in X', \quad h_n \in X, \text{ and } h_{n-1} = h_n', \\ \qquad \text{if there is such an n}, \\ \infty, \quad \text{otherwise.} \end{cases}$$

One verifies by imitating the proof of Theorem 5.1 that the terminal gamble $\gamma \in \Gamma^{ran}(x_0)$ can be realized by using the modified strategy σ^* and the modified stopping time t^*. Unfortunately, the new σ^* is not available in the old house Γ^l.

Acknowledgement

We are indebted to Eugene Feinberg and the referee for bringing to our attention the result of Dynkin and Yushkevich cited above and several other references in the literature on Markov decision processes.

References

[1] R.J. Aumann, Mixed and behavior strategies in infinite extensive games, *Ann. Math. Studies* 53(1964), 627–650.

[2] R.J. Aumann, Borel structures for function spaces, *Illinois J. Math.* 5(1961), 614-630.

[3] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control : The Discrete Time Case*, Academic Press, New York, 1978.

[4] C. Dellacherie, Quelques résultats sur les maisons de jeux analytiques, *Séminaire de Probabilités XIX, Strasbourg 1983–84*, Lecture Notes in Math., vol. 1123, Springer-Verlag, Berlin and New York, 1985, pp. 222–229.

[5] C. Dellacherie and P. A. Meyer, *Probabilitiés et Potentiel*, Hermann, Paris, Chapters IX–XI, 1983.

[6] L.E. Dubins and L.J. Savage, *Inequalities for Stochastic Processes*, Dover, New York, 1976.

[7] E.B. Dynkin and A.A. Yushkevich, *Controlled Markov Processes*, Springer-Verlag, Berlin, 1979.

[8] E.A. Feinberg, Nonrandomized Markov and semi-Markov strategies in dynamic programming, *SIAM Theory of Prob. and Appl.* 27 (1982), 116-126.

[9] E.A. Feinberg, Controlled Markov processes with arbitrary numerical criteria, *SIAM Theory of Prob. and Appl.* 27 (1982a), 486-503.

[10] E.A. Feinberg, Sufficient classes of strategies in discrete dynamic programming. I: decomposition of randomized strategies and imbedded models, *SIAM Theory of Prob. and Appl.* 31 (1986), 658-668.

[11] E.A. Feinberg, Nonrandomized strategies in stochastic decision processes, *Ann. Oper. Res* 29 (1991), 315-332.

[12] E. A. Feinberg, On measurability and representation of strategic measures in Markov decision processes, in : T. S. Ferguson, L. S. Shapley and J. B. MacQueen (editors), *Statistics, Probability and Game Theory,* Lecture Notes-Monograph Series, vol. 30, Institute of Math. Statist., Hayward, CA, 1996, pp. 29–43.

[13] I.I. Gikhman and A.V. Skorohod, *Controlled Random Processes*, Springer, New York, 1979.

[14] T.P. Hill and V.C. Pestien, The existence of good Markov strategies for decision processes with general payoffs, *Stoch. Proc. and Appl.* 24 (1987), 61-76.

[15] A. S. Kechris, *Classical Descriptive Set Theory,* Springer-Verlag, Berlin and New York, 1995.

[16] N.V. Krylov, Construction of an optimal strategy for a finite controlled chain, *SIAM Theory of Prob. and Appl.* 10 (1965) 45-54.

[17] H.W.Kuhn, Extensive games and the problem of information, *Ann. Math. Studies* 28 (1953), 193-216.

[18] A. Maitra, R. Purves and W. Sudderth, Leavable gambling problems with unbounded utilities,*Trans. Amer. Math. Soc.* 320 (1990), 543–567.

[19] A. Maitra and W. Sudderth, Saturations of gambling houses, *Séminaire de Probabilités*, to appear.

[20] B.V.Rao, Borel structures for function spaces, *Colloquium Mathematicum* 23 (1971), 33-38.

School of Statistics, University of Minnesota, Minneapolis, Minnesota 55455.