NONEXISTENCE OF MEASURABLE OPTIMAL SELECTIONS

JOHN BURGESS AND ASHOK MAITRA

(Communicated by Lawrence F. Gray)

Abstract. We give an example of a function $f$ on a separable metric space $X$ into a compact metric space $Y$ such that the graph of $f$ is a Borel subset of $X \times Y$, but $f$ is not Borel measurable. The example forms the basis for our construction of an upper semicontinuous, compact model of a one-day dynamic programming problem where the player has an optimal action at each state, but is unable to make a choice of such an action in a Borel measurable manner.

1. Introduction

A very useful result for dynamic programming/gambling/discrete time stochastic control is the following selection theorem due to Schäl [8, 9]:

Theorem A. Suppose $(T, \mathcal{T})$ is a measurable space, and let $Y$ be a separable metric space. Suppose $G$ is a $\mathcal{T}$-measurable multifunction on $T$ such that $G(t)$ is a nonempty, compact subset of $Y$ for every $t \in T$. Let $v$ be a real-valued function on $\text{Graph}(G) (= \{(t, y) \in T \times Y : y \in G(t)\})$ satisfying the following condition:

(1) $v$ is the pointwise limit of a nonincreasing sequence $\{v_n\}$ of real-valued functions on $\text{Graph}(G)$ such that each $v_n$ is $(\mathcal{T} \times \mathcal{B}_Y|\text{Graph}(G))$-measurable and $v_n(t, \cdot)$ is continuous on $G(t)$ for every $t \in T$, where $\mathcal{B}_Y$ is the Borel $\sigma$-field on $Y$. Let

$$v^*(t) = \sup_{y \in G(t)} v(t, y), \quad t \in T.$$  

Then there is a $\mathcal{T}$-measurable selector $g: T \rightarrow Y$ of $G$ such that $v^*(t) = v(t, g(t))$ for every $t \in T$.

We recall that to say $G$ is $\mathcal{T}$-measurable means that for every open subset $V$ of $Y$, the set $\{t \in T : G(t) \cap V \neq \emptyset\}$ belongs to $\mathcal{T}$. Since $G$ is compact valued, this condition implies the apparently stronger condition that for every closed subset $C$ of $Y$, the set $\{t \in T : G(t) \cap C \neq \emptyset\}$ belongs to $\mathcal{T}$.

In the dynamic programming literature, theorems of the above type are called Dubins-Savage selection theorems. Theorem A is the culmination of many attempts (see, for instance [4, 11]) to improve on the original result of Dubins and Savage [1]. Many infinite horizon problems of dynamic programming can
be reduced to a one-day problem by use of the Optimality Equation. Once this is done, the situation becomes like that described in Theorem A. One can think of $T$ as the state space, $G(t)$ as the set of actions available to the player when the system is in the state $t$, and $v(t, y)$ as the reward when the player chooses action $y$ in state $t$. Theorem A then states that the player can choose an optimal action in each state in a measurable manner. This optimal rule can then be used every day to define an optimal stationary policy in the infinite horizon problem (see, e.g., [4, 9] for details).

It is not unreasonable to conjecture that Theorem A remains true if condition (1) is replaced by the weaker condition

$$(1') \quad v \text{ is } (\mathcal{F} \times \mathcal{B}_p|\text{Graph}(G))-\text{measurable and } v(t, \cdot) \text{ is upper semicontinuous (u.s.c.) on } G(t) \text{ for every } t \in T.$$  

The aim of this article is to give a counterexample to this conjecture, even when $T$ is a separable metric space and $\mathcal{F}$ is its Borel $\sigma$-field.

2. Coding of Borel sets

Our basic example will be based on a coding of Borel subsets of $\mathbb{R}$, the real line, due to Solovay [10, pp. 25–28]. The coding will now be described in some detail. The main properties of the coding will be stated in a sequence of lemmas. The proofs of the lemmas can be found in [10].

In what follows, $\omega$ will denote the set of natural numbers. Let $\{r_i\}$ be an enumeration of the rationals and let $J$ be the pairing function on $\omega \times \omega$ defined by $J(a, b) = 2^a(2b + 1)$. We define the coding recursively as follows:

1. $a \in \omega^\omega$ codes $[r_i, r_j]$ if $a(0) = 0 \pmod{3}$, $a(1) = i$, and $a(2) = j$.
2. Suppose $a_i \in \omega^\omega$ codes $B_i \subseteq \mathbb{R}$, $i = 0, 1, \ldots$; then $a \in \omega^\omega$ codes $U_i B_i$ if $a(0) \equiv 1 \pmod{3}$ and $a(J(a, b)) = a_\alpha(b)$.
3. Suppose $\beta \in \omega^\omega$ codes $B$, $\alpha(0) = 2 \pmod{3}$, and $\alpha(n + 1) = \beta(n)$. Then $a$ codes $B^c$, the complement of $B$.
4. $a \in \omega^\omega$ codes $\mathbb{R}$ only as required by 1–3.

Lemma 1. (a) Every $a \in \omega^\omega$ codes at most one subset of $\mathbb{R}$.

(b) Every Borel subset of $\mathbb{R}$ is coded by some $a \in \omega^\omega$.

(c) If a subset of $\mathbb{R}$ is coded by $a$, then it is Borel.

Next, we define a function $\Phi: \omega^\omega \times \omega \to \omega^\omega$, with the property that if $a$ codes a Borel set $B$, then $\Phi(a, \cdot)$ recovers the Borel sets from which $B$ is constructed. For the definition, we need to fix an enumeration $\{s_n\}$ without repetitions, of finite sequences of natural numbers (including the empty sequence) such that if $s_n$ is an initial segment of $s_m$, then $n \leq m$. So $s_0$ is the empty sequence. The definition of $\Phi(a, n)$ will proceed by induction on $n$.

Set $\Phi(a, 0) = a$, $a \in \omega^\omega$. Let $n > 0$ and suppose that $\Phi(a, m)$ has been defined for all $a \in \omega^\omega$ and all $m < n$. Since $n > 0$, $s_n$ is a finite sequence of positive length $k$, say. Let $s_m$ be the initial segment of $s_n$ of length $(k - 1)$, so $m < n$; let $u$ be the last coordinate of $s_n$. Now define for $i \in \omega$

$$\Phi(a, n)(i) = \begin{cases} 0 & \text{if } \Phi(a, m)(0) \equiv 0 \pmod{3}, \\ \Phi(a, m)(J(u, i)) & \text{if } \Phi(a, m)(0) \equiv 1 \pmod{3}, \\ \Phi(a, m)(i + 1) & \text{if } \Phi(a, m)(0) \equiv 2 \pmod{3}. \end{cases}$$
Lemma 2. (a) $\Phi$ is Borel measurable.
(b) If $\alpha$ codes a Borel set, then for all $n$, $\Phi(\alpha, n)$ codes a Borel set.

For $\beta \in \omega^\omega$, define $\overline{\beta} \in \omega^\omega$ so that for every $n \in \omega$,
$$s_{\overline{\beta}(n)} = (\beta(0), \beta(1), \ldots, \beta(n-1)).$$
Plainly, the map $\beta \rightarrow \overline{\beta}$ is continuous.

Let $C = \{\alpha \in \omega^\omega : (\forall \beta)(\exists n)\Phi(\alpha, \overline{\beta}(n)) \equiv 0\}.$

Lemma 3. (a) $C$ is a coanalytic subset of $\omega^\omega$.
(b) $C$ is the set of all $\alpha \in \omega^\omega$ such that $\alpha$ codes a Borel subset of $\mathbb{R}$.

Define a function $\varphi : \omega \times \omega \rightarrow \omega$ such that $s_{\varphi(n, i)}$ is the sequence $s_n$ followed by $i$.

Let $E \subseteq \omega^\omega \times \mathbb{R} \times 2^\omega$ be defined as follows:
$$(\alpha, x, y) \in E \iff (\forall n)[(\Phi(\alpha, n)(0) \equiv 0 \pmod{3})$$
$$\rightarrow \{\gamma(n) = 1 \iff (\exists i)(\exists j)(\Phi(\alpha, n)(1) = i \& \Phi(\alpha, n)(2) = j \& x \in [r_i, r_j])\}]$$
$$\& (\forall n)[(\Phi(\alpha, n)(0) \equiv 1 \pmod{3} \rightarrow \{\gamma(n) = 1 \iff (\exists i)(\Phi(\alpha, n)(1) = i)\}]$$
$$\& (\forall n)[(\Phi(\alpha, n)(0) \equiv 2 \pmod{3} \rightarrow \{\gamma(n) = 1 \iff \Phi(\alpha, n)(0) = 0\}]$$

Lemma 4. (a) $E$ is a Borel subset of $\omega^\omega \times \mathbb{R} \times 2^\omega$.
(b) For every $\alpha \in C$ and $x \in \mathbb{R}$, there is a unique $\gamma \in 2^\omega$ such that $(\alpha, x, \gamma) \in E$.
(c) If $\alpha \in C$ and $(\alpha, x, \gamma) \in E$, then $(\forall n)(\gamma(n) = 1 \iff x \in \text{the Borel set coded by } \Phi(\alpha, n)).$

This completes our description of Solovay’s coding of Borel subsets of $\mathbb{R}$.

3. COUNTEREXAMPLES

We use the notation introduced in the previous section.

Let $f$ be the function on $C \times \mathbb{R}$ into $2^\omega$ guaranteed by Lemma 4(b); that is, for $\alpha \in C$ and $x \in \mathbb{R}$, $f(\alpha, x)$ is the unique $\gamma$ such that $(\alpha, x, \gamma) \in E$.

It follows that
$$\text{Graph}(f) = E \cap (C \times \mathbb{R} \times 2^\omega)$$
so that $\text{Graph}(f)$ is (relatively) Borel in $C \times \mathbb{R} \times 2^\omega$. We will now show that $f$ is not Borel measurable. Towards a contradiction, assume that $f$ is Borel measurable on $C \times \mathbb{R}$.

Consider the set
$$F = \{ (\alpha, x) \in C \times \mathbb{R} : f(\alpha, x)(0) = 0 \}.$$

According to Lemmas 3(b) and 4(c) and the definition of $\Phi$, the condition "$\alpha \in C \& f(\alpha, x)(0) = 0"$ states that $x$ does not belong to the Borel set coded by $\alpha$. Since $f$ is Borel measurable, $F$ is (relatively) Borel in $C \times \mathbb{R}$ so there must exist a Borel subset $D$ of $\omega^\omega \times \mathbb{R}$ such that $F = D \cap (C \times \mathbb{R})$.

Fix a homeomorphism $h$ from the set $I$ of irrationals in $\mathbb{R}$ onto $\omega^\omega$. Let $B = \{ x \in I : (h(x), x) \in D \}$. Plainly, $B$ is a Borel subset of $\mathbb{R}$. So, by Lemmas 1(b) and 3(b), there is $\alpha^* \in C$ such that $\alpha^*$ codes $B$. Set $x^* = h^{-1}(\alpha^*)$. Then
$$x^* \in B \iff (\alpha^*, x^*) \in D \iff (\alpha^*, x^*) \in F \iff f(\alpha^*, x^*)(0) = 0$$
$$\iff x^* \notin B,$$
a contradiction!
Since the domain of the function \( f \) is coanalytic (Lemma 3(a)) and \( 2^\omega \) is homeomorphic to the Cantor set, the above construction shows that the Measurable Graph Theorem fails, in general, for a function with coanalytic domain \( X \) and values in \([0, 1]\). That is, such a function can have a graph that is a Borel subset of \( X \times [0, 1] \), but it is not necessarily the case that the function is Borel measurable. A fortiori, a multifunction with coanalytic domain and nonempty, compact subsets of \([0, 1]\) as values can have a Borel measurable graph without being Borel measurable. This answers a question of some interest to workers in stochastic optimization (see, e.g., [6]). On the other hand, it is a classical result (see [3]) that Borel measurability of the graph of a function with analytic domain and values in a Polish space implies the Borel measurability of the function. This is also true, courtesy of the Kunugui-Novikov theorem [3], of a multifunction with analytic domain and nonempty, compact values in a Polish space.

It is now easy to turn the basic example into the desired counterexample to Theorem A with condition (1) replaced by (1').

Let \( T = C \times \mathbb{R} \), \( \mathcal{T} \) the Borel \( \sigma \)-field of \( C \times \mathbb{R} \), \( Y = 2^\omega \), \( G(t) = 2^\omega \) for every \( t \in T \), and

\[
v((\alpha, x), y) = \begin{cases} 1 & \text{if } f(\alpha, x) = y, \\ 0 & \text{otherwise}. \end{cases}
\]

Then \( v \) is Borel measurable on \( Graph(G) = T \times Y \), since \( Graph(f) \) is a Borel subset of \( T \times Y \). Moreover for fixed \( t \in T \), \( v(t, \cdot) \) is u.s.c. on \( G(t) \). Thus, condition (1') is satisfied. But, plainly, there is no Borel measurable selector \( g \) of \( G \) such that \( v^*(t) = v(t, g(t)) \) for all \( t \in T \).

In the example above, the function \( v^* \), being identically equal to one, is \( \mathcal{F} \)-measurable. We now modify the example so that \( v^* \) is not \( \mathcal{F} \)-measurable.

Fix a Borel isomorphism of the Borel set \( E \) and the interval \([1/2, 1]\) and denote by \( \psi \) the restriction of the Borel isomorphism to \( Graph(f) \), so that \( \psi \) continues to be a Borel isomorphism of \( Graph(f) \) and the range of \( \psi \). As before, take \( T = C \times \mathbb{R} \), \( \mathcal{T} \) the Borel \( \sigma \)-field of \( C \times \mathbb{R} \), \( Y = 2^\omega \), \( G(t) = 2^\omega \) for every \( t \in T \), and

\[
v((\alpha, x), y) = \begin{cases} \psi(\alpha, x, y) & \text{if } f(\alpha, x) = y, \\ 0 & \text{otherwise}. \end{cases}
\]

It is easily checked that condition (1') is satisfied. Next observe that

\[
v^*((\alpha, x)) = \psi(\alpha, x, f(\alpha, x))
\]

for all \((\alpha, x) \in T\). If \( v^* \) were \( \mathcal{F} \)-measurable, then the map \((\alpha, x) \to (\alpha, x, f(\alpha, x))\) would be \( \mathcal{F} \)-measurable, since \( \psi \) is a Borel isomorphism. But then \( f \) would be \( \mathcal{F} \)-measurable, a contradiction!

We remark that if \( T \) is an analytic set and \( \mathcal{T} \) the Borel \( \sigma \)-field of \( T \), then condition (1') implies condition (1) and so Theorem A is true under condition (1'). To see that (1') implies (1), fix a metric \( \rho \) on \( Y \) and define for \( n = 1, 2, \ldots \),

\[
v_n(t, y) = \sup_{y' \in G(t)} \{v(t, y') - n\rho(y', y)\}, \quad (t, y) \in Graph(G).
\]

The only nontrivial thing to check regarding condition (1) is that \( v_n \) is \((\mathcal{T} \times BY|Graph(G))\)-measurable and this follows from the Kunugui-Novikov theorem.
The validity of Theorem A under condition \((1')\) when \(T\) is analytic and \(\mathcal{F}\) is its Borel \(\sigma\)-field was observed by Himmelberg et al. [2] and also by Schäl [11].

4. Proof of Theorem A

For the sake of completeness, we give a quick proof of Theorem A. Our proof differs from those of Schäl [8, 9] and Reider [7]. We start with an easy lemma.

**Lemma B.** Let \(Z\) be a compact metric space. Suppose \(f, f_m : Z \to \mathbb{R}\) are u.s.c. and \(f_m \downarrow f\). If \(z_m \to z_0\) in \(Z\), then \(\lim_{m} \sup_{m} f_m(z_m) \leq f(z_0)\).

**Proof.** Let \(\varepsilon > 0\). Choose a continuous function \(g : Z \to \mathbb{R}\) such that \(g \leq f\) and \(f(z_0) \geq g(z_0) - \varepsilon\).

Set \(g_m = \max(f_m, g), \ m \geq 1\). Then \(g_m\) is u.s.c. and \(g_m \downarrow g\). Dini's theorem now applies to yield that the convergence \(g_m \downarrow g\) is uniform on \(Z\). Hence

\[
\lim_{m} \sup_{m} f_m(g_m) \leq \lim_{m} \sup_{m} g_m(z_m) = g(z_0) \leq f(z_0) + \varepsilon.
\]

**Proof of Theorem A.** By [5, Theorem 5], for each \(n \geq 1\), there is a \(\mathcal{F}\)-measurable selector \(g_n : T \to Y\) of \(G\) such that

\[
v_n(t, g_n(t)) = \sup_{y \in G(t)} v_n(t, y)
\]

for every \(t \in T\).

For \(t \in T\) set \(H(t) = \{y \in G(t) : \text{there is a subsequence } \{g_{n_i}(t)\} \text{ such that } g_{n_i}(t) \to y\}\). Plainly, \(H(t)\) is nonempty, compact as \(G(t)\) is. We claim that \(H\) is \(\mathcal{F}\)-measurable. Let \(C\) be closed in \(Y\). Then

\[
\{t \in T : H(t) \cap C \neq \emptyset\} = \bigcap_{k \geq 1} \bigcup_{m \geq k} \left\{t \in T : \rho(g_m(t), C) < \frac{1}{k}\right\},
\]

where \(\rho\) is a metric on \(Y\). Since the set on the right belongs to \(\mathcal{F}\), \(H\) is \(\mathcal{F}\)-measurable. Use the Kuratowski-Ryll-Nardzewski selection theorem [3] to get a \(\mathcal{F}\)-measurable selector \(g : T \to Y\) of \(H\).

To complete the proof, fix \(t \in T\) and let \(y^* = g(t)\). So there is a subsequence \(\{g_{n_i}(t)\}\) such that \(y_{n_i} = g_{n_i}(t) \to y^*\). By condition (1) and Lemma B, we have

\[
\lim_{i} v_{n_i}(t, y_{n_i}) \leq v(t, y^*).
\]

It follows that \(v^*(t) = v(t, g(t))\).

**References**


**Department of Philosophy, Princeton University, Princeton, New Jersey 08544-1006**

**Department of Mathematics, Santa Fe Institute, Santa Fe, New Mexico 87501**