A Bayesian Game without ϵ-equilibria

Robert Samuel Simon and Grzegorz Tomkowicz

July 7, 2017

Robert Simon
London School of Economics
Department of Mathematics
Houghton Street
London WC2A 2AE
email: R.S.Simon@lse.ac.uk

Grzegorz Tomkowicz
Centrum Edukacji G^2
ul.Moniuszki 9
41-902 Bytom
Poland
e-mail: gtomko@vp.pl

Abstract

We present a three player Bayesian game for which there are no ϵ-equilibria in Borel measurable strategies for small enough positive ϵ, however there are non-measurable equilibria. The structure of the game employs a non-amenable semi-group action corresponding to the knowledge of the players. The equilibrium property is related to the proper colouring of graphs and the Borel chromatic number; but rather than keeping adjacent vertices coloured differently there are algebraic conditions relating to the topology of the space and some ergodic operators.

Key words: Bayesian games, non-amenable semi-group action, equilibrium existence

1 Introduction

A fundamental concept of game theory is that of an equilibrium, a determination of strategies for each player such that no player prefers to switch to another strategy, given that the strategies of the other players remain fixed. An ϵ-equilibrium for any $\epsilon \geq 0$ is defined in the same way – no player prefers by more than ϵ to switch to another strategy. With any class of games, one asks if all games of the class possess equilibria. If not, does every game of the class at least possess an ϵ-equilibrium for every positive ϵ?

A Bayesian game is a game of incomplete information, a game for which players have private information. Bayesian games are ancient, most card games being good examples. What happens when the possibilities for private information become uncountable in number? In that context, what do we mean by a player preferring some strategy over another by at least ϵ?

To focus on the complexities of the information structure, we assume throughout that there are finitely many players and to each player finitely many actions.

The conventional assumption is that the payoff from a strategy in a Bayesian game is an integration of a measurable function over the whole space of possibilities. Usually we assume that there is a topological space Ω with a probability measure μ defined on the collection $\mathcal{F}$ of Borel sets and that each player i is assigned a sigma algebra $\mathcal{F}_i$ contained in $\mathcal{F}$. The sigma algebra $\mathcal{F}_i$ corresponds to player i's knowledge of the space and also defines that player's strategy space as the set of functions from Ω to the set of probability distributions on that player's actions which are measurable with respect to $\mathcal{F}_i$. We assume also that for every choice of actions, one for each player, the payoff to each player is $\mathcal{F}$ measurable (or even continuous). In this way, a collection of strategies, one for each player, defines an expected payoff for each of the players. If there is an equilibrium in this context, it is called a *Harsanyi* equilibrium (Simon 2003).

It was demonstrated (Simon 2003) that there is a three player game defined in this way that lacks a Harsanyi equilibrium, yet does possess a different kind of equilibrium when the players are allowed strategies that are not measurable with respect to their Borel fields and their payoff evaluations are local in character and do not involve integrating over the whole space Ω. Equilibria of this latter kind are called *Bayesian equilibria* (Simon 2003).

The definition of a Bayesian equilibrium requires additional conditions on the structure of the game. In particular, for each player i the sigma algebra $\mathcal{F}_i$ has to be generated by a partition $\mathcal{P}_i$ of Ω so that $\mathcal{F}_i$ is the collection of sets such that $B \in \mathcal{F}_i$ if and only if $B \in \mathcal{F}$ and for every $A \in \mathcal{P}_i$ either $A \subseteq B$ or $A \cap B = \emptyset$. Furthermore, for every $x \in A \in \mathcal{P}_i$ there is a regular conditional probability with respect to μ with support in A that is constant inside of A, and furthermore there is a countable or finite subset $A' \subseteq A$ such that the payoff is 0 to player i at all points in $A \backslash A'$ no matter what actions are chosen by the players. A strategy for player i is now allowed to be any function from Ω to the probability distributions on that player's actions such that for every $A \in \mathcal{P}_i$ the function is a constant inside of A. In this way, using the regular conditional probabilities, a local expected payoff is defined for player i in all members A of $\mathcal{P}_i$ regardless of the strategies chosen. A Bayesian ϵ-equilibrium is defined to be a collection of such strategies, one for each player, so that for every player i and every set $A \in \mathcal{P}_i$ the player i cannot obtain an improvement of more than ϵ in his or her expected payoff relative to that set $A \in \mathcal{P}_i$ by choosing a different strategy.

While Harsanyi equilibria do not always exist, the question remained, however, does there exist always Harsanyi ϵ-equilibria?

We answer this question in the negative; there is a three-player Bayesian game where each player has two actions for which there are no Harsanyi ϵ-equilibria for all $\epsilon \leq \frac{1}{1000}$, and yet, assuming the Axiom of Choice, there are non-measurable Bayesian equilibria that employ pure strategies almost everywhere (pure meaning that the strategies map to the two extremal points of the one-dimensional probability simplex).

Harsanyi (1967) introduced a global theoretical perspective to Bayesian games. Milgrom and Weber (1985) asked implicitly the question whether Bayesian games always have measurable equilibria after proving existence for a special class of Bayesian games and analysing a game which did not belong to that class yet still had Borel measurable equilibria.

A significant advance was performed by Hellman (2014). He showed that there is a two-player Bayesian game with Bayesian equilibria but no Bayesian ϵ-equilibrium that is also Borel measurable for small enough positive ϵ. This discovery was advanced further by Hellman and Levy (2016), who demonstrated that a broad class of knowledge structures support games for which the same holds. The Hellman and Levy paper serves well as a general source to the structure, problems, and history of Bayesian games, especially in their relation to countable equivalence relations that are amenable.

What is the relationship between a Harsanyi ϵ-equilibrium and a Borel measurable Bayesian ϵ-equilibrium? It is required of a Bayesian ϵ-equilibrium that throughout the space each player cannot gain locally more than ϵ through an alternative strategy. However in a Harsanyi ϵ-equilibrium there could be a player who, according to the strategies defining the ϵ-equilibrium, could improve his or her payoff by as much as $B > 0$ at a set of measure no more than ϵ/B. A Borel measurable Bayesian ϵ-equilibrium is a Harsanyi ϵ-equilibrium, but not necessarily vice versa. Indeed it is not difficult to show that for every $\epsilon > 0$ there are Harsanyi ϵ-equilibria to the Hellman game cited above.

One can perceive sets of very small measure where a player can act foolishly as a kind of firewall, absorbing the conflicts between the measurability and equilibrium requirements. Amenable structures tend to allow for such firewalls; for example with the related topic of Borel colouring; see Kechris, et all (1999). Therefore we would not have expected to find a game example lacking Harsanyi ϵ-equilibria (yet possessing Bayesian equilibria) without utilising a non-amenable structure to the knowledge of the players.

As long as a game has an Harsanyi ϵ-equilibrium for every positive ϵ there is an equilibrium payoff, namely a cluster point of payoffs corresponding to the ϵ-equilibria as ϵ goes to 0. By this interpretation of an equilibrium payoff, ours is a Bayesian game that has equilibria, but no equilibrium payoff.

With our example, there is no proper subset of the probability space for which the players have common knowledge, hence the arguments used are different from that of previous Bayesian games that lack Harsanyi equilibria but have Bayesian equilibria, which do utilise countable equivalence relations. Rather there is a directional relation of influence such that behaviour at every point is influenced by what happens at at most countably many other points.

After the first two lemmas, the game theoretic aspects are replaced by a concern with the "parity rule", an algebraic condition on functions from a probability space, acted upon by a free semi-group with two generators, to the group of order two. Although strictly speaking the parity rule is different from the proper colouring of a graph, the relation to such colouring is unmistakable. We believe that our result will inspire work in related directions involving similar algebraic conditions, with or without any further connections to game theory. Further interest here includes the measurability of the functions with respect to finitely additive extensions of the given probability measure.

The rest of this paper is organised as follows. In the next section we define the game, followed in the third section by a proof that this game has no Harsanyi $\frac{1}{1000}$-equilibrium. In the fourth section we show that it does have non-measurable Bayesian equilibria. The concluding fifth section is a presentation of open problems.

2 The Game:

Let $G^+ = \mathbf{F}_2^+$ be the free semi-group generated by the non-negative powers of two independent generators T_1 and T_2, with e the identity included in G^+. Let X be the space $\{0,1\}^{G^+}$, with an $x \in X$ a collection of the form $(x^e, x^{T_1}, x^{T_2}, x^{T_1T_2}, x^{T_2T_1}, x^{T_1^2}, x^{T_2^2}, \dots)$ with $x^U \in \{0,1\}$ for every $U \in G^+$. A cylinder of X is determined by a finite subset $\mathcal{U}$ of G^+ and a specified value in $\{0,1\}$ for each of the $U \in \mathcal{U}$ (with either $x^U = 0$ or $x^U = 1$ allowed for all $U \notin \mathcal{U}$). An open set of X is an arbitrary union of cylinders.

For both $i = 1,2$ define $T_i : X \to X$ to be the shift: $T_i(x)^V = x^{T_iV}$ for all $V \in G^+$, and for every $U, V \in G^+$ define $U : X \to X$ by $U(x)^V = x^{UV}$. This defines a right action of G^+ on X, meaning that $UV(x) = V \circ U(x)$. For every $x \in X$ define $G^+(x)$ to be the countable set $\{U(x) \mid U \in G^+\}$. Very important to the structure of X is that for every $x, y \in X$ there are two $z_0, z_1 \in X$ such that $T_1(z_i) = x$ and $T_2(z_i) = y$ for both $i = 0, 1$, and they differ only in the e location. We call the points z_0, z_1 the *twins determined* by x and y. We place the canonical probability distribution m on X which gives $\frac{1}{2}$ to each 0 or 1 placed in each position of G^+ and independently, so that a cylinder defined by n positions is given the probability 2^{-n}. With this probability distribution, we see that all $U \in G^+$ are measure preserving actions, meaning that $m(U^{-1}A) = m(A)$ for all Borel subsets A.

Of special importance, the independent probability on each position implies that the distribution m on X can be reconstructed from the measure preserving property of the $T_1 : X \to X$ and $T_2 : X \to X$, its product distribution m^2, combined with an equal probability given to both twins.

Let D be the set $D := \{r, g\}$, r for red and g for green. The probability space on which the game is played will be $\Omega := D \times X$. We define the topology on

Ω to be that induced by the clopen (closed and open) sets defined by the set D and the cylinders of X, so that Ω is a Cantor set. We define the canonical probability distribution μ on Ω so that for each choice of $d \in D$ and 0 or 1 in n distinct positions the probability for this cylinder will be $\frac{1}{2}\ 2^{-n}$. For example, μ gives the set $\{(r,x) \mid x^e = 0, x^{T_1} = 1\}$ the probability $\frac{1}{8}$. The measure μ is the *common prior* for the game, meaning the Borel probability measure by which the game is defined.

There are three players, labelled G_0, R_1, R_2. The *information sets* of a player are the sets that partition the space Ω from which that player's Borel field is defined in the way described above. What we call the *belief* of a player at one of his or her information sets is the regular conditional probability that is constant within that set. The information sets of each player are defined as follows. For each $x \in X$ Player G_0 considers (g,x) and (r,x) possible, with a belief in both points of equal $\frac{1}{2}$ probability, and these two points constitute its information set. For each $i = 1,2$ and each x Player R_i's considers (r,x) and $\{g\} \times T_i^{-1}(x)$ possible, and this pairing of a point with the corresponding Cantor set is its information set. Player R_i's belief is that the point (r,x) and the set $\{g\} \times T_i^{-1}x$ are equally likely, with $\frac{1}{2}$ probability given to both. Notice that this belief by the player R_i is consistent with the probability distribution μ on Ω, as the measure preserving property of the T_i implies that $m(T_i^{-1}(A)) = m(A)$ for all Borel subsets A of X. Within the set $\{g\} \times T_i^{-1}x$ Player R_i's belief is the canonical one consistent with the measure μ, giving all cylinders of the same length the same probability. If B is the information set of a player, it means that this player cannot distinguish between any two points of this set and therefore has to conduct the same behaviour throughout the set. Otherwise, as described, we place no further restrictions on the strategies of the players.

All players have only two actions. The red players R_1 and R_2 have the actions $\mathbf{a_0}, \mathbf{a_1}$ and the green player G_0 has the actions $\mathbf{b_0}$, $\mathbf{b_1}$.

For either player R_i the only payoff that matters is that obtained at those states labelled r, and for the player G_0 the same is true for those states labelled g. There are two equivalent approaches to be taken, illustrated for a player R_i. Either the payoff obtained at (r,x), described below, is duplicated at all the other points in the same information set, namely the set $\{g\} \times T_i^{-1}(x)$, or the payoffs obtained at (r,x), described below, is multiplied by 2 and at all other points in the same information set the payoff is 0.

Though the latter interpretation may be better suited to some theoretical approaches, as it employs the probability distribution μ on Ω, we will assume throughout the former equivalent interpretation (and for Player G_0 as well at the information sets of the form $\{(r,x),(g,x)\}$). This will allow us to focus on the set X and its probability distribution m.

The $\mathbf{b_0}$ and $\mathbf{b_1}$ pertain to actions of Player G_0 at both (g,x) and (r,x). If $x^e = 0$ then the payoff matrices for the players R_i at the states (r,x) are

R_1		$\mathbf{b_0}$	$\mathbf{b_1}$		R_2		$\mathbf{b_0}$	$\mathbf{b_1}$
	$\mathbf{a_0}$	300	0	and		$\mathbf{a_0}$	100	0 .
	$\mathbf{a_1}$	0	100			$\mathbf{a_1}$	0	300

If $x^e = 1$ then the payoff matrices at (r,x) are reversed:

R_1		$\mathbf{b_0}$	$\mathbf{b_1}$		R_2		$\mathbf{b_0}$	$\mathbf{b_1}$
	$\mathbf{a_0}$	100	0	and		$\mathbf{a_0}$	300	0 .
	$\mathbf{a_1}$	0	300			$\mathbf{a_1}$	0	100

More complex are the payoffs of the player G_0 at a state (g,x). The matrix is three dimensional, meaning that it is a $2 \times 2 \times 2$ matrix. We need only to describe a 2×2 matrix corresponding to each action of the G_0 player. The rows and columns stand for the actions of the R_1 and R_2 players, respectively. Those actions $\mathbf{a_0}$ and $\mathbf{a_1}$ are performed by the R_1 player at both (g,x) and $(r, T_1 x)$ and by the R_2 player at both (g,x) and $(r, T_2 x)$. First we describe the payoff matrices if $x^e = 0$:

$\mathbf{b_0}$		$\mathbf{a_0}$	$\mathbf{a_1}$	$\mathbf{b_1}$		$\mathbf{a_0}$	$\mathbf{a_1}$
	$\mathbf{a_0}$	1000	0		$\mathbf{a_0}$	0	1000
	$\mathbf{a_1}$	0	2000		$\mathbf{a_1}$	2000	0

On the other hand, if $x^e = 1$ then the structure of payoffs is reversed:

$\mathbf{b_0}$		$\mathbf{a_0}$	$\mathbf{a_1}$	$\mathbf{b_1}$		$\mathbf{a_0}$	$\mathbf{a_1}$
	$\mathbf{a_0}$	0	1000		$\mathbf{a_0}$	2000	0
	$\mathbf{a_1}$	2000	0		$\mathbf{a_1}$	0	1000

A strategy of a player is a function from its collection of information sets to the probability distributions on its two actions (a one dimensional simplex). The strategy is Borel measurable if that function is measurable with respect to its Borel sigma algebra (which is defined canonically as above from its information sets).

However the G_0 player acts at some (g, x), that action is copied at (r, x) (because the G_0 player cannot distinguish between these two points). However the R_i players respond at (r, x), those actions are copied at the sets $\{g\} \times T_i^{-1}(x)$ respectively (as the R_i player cannot distinguish between (r, x) and $\{g\} \times T_i^{-1}(x)$). The behaviour of a player at (g, x) or (r, x) will influence inductively the behaviour of all players at an uncountable subset leading upward through repetitive applications of the T_i^{-1}. In the other direction, the behaviour that influences inductively a player's payoff at (g, x) or (r, x) lies entirely within the countable set $D \times G^+(x)$. With regard to this latter aspect of influence (rather than influencing), our game shares similarity with those defined by countable equivalence relations.

3 No Harsanyi $\frac{1}{1000}$-equilibria

Before we show that the game has no Harsanyi $\frac{1}{1000}$-equilibrium, we focus on the subset $\{g\} \times X$.

Let A_0 be the subset of X such that Player G_0 at $\{g\} \times A_0$ chooses $\mathbf{b_0}$ with probability at least $\frac{19}{20}$. Let A_1 be the corresponding subset of X such that Player G_0 chooses $\mathbf{b_1}$ with probability at least $\frac{19}{20}$. Let A_M be the subset $X \backslash (A_0 \cup A_1)$.

As a general rule, from the above payoff matrices and the assumption that players are following their interests (the interests of the R_i players at (r, x) being that of conveying the choice of the G_0 player at (g, x)), we would expect that if $T_1(x) \in A_i$ and $T_2(x) \in A_j$, and $x^e = k$ then $x \in A_{i+j+k}$ where $i+j+k$ is represented modulo two. We call this the *parity rule*, and say that this rule holds for a point x whenever these three belongings are true.

If any player chooses both actions at some point with strictly more than $\frac{1}{20}$ we say that the player is *mixing* at that point (meaning $x \in A_M$ when this player is G_0). If there is a player and a set A of measure at least $w > 0$ where that player prefers one strategy over another by at least $r > 0$ and either that player is mixing or choosing the non-preferred action by at least $\frac{1}{20}$, then that player can gain at least $\frac{rw}{20}$ by choosing a different strategy. Therefore in an ϵ-equilibrium it follows that w is at most $\frac{20\epsilon}{r}$. In particular, we use in many places $r = 80$ and $\epsilon = \frac{1}{1000}$, hence $w \leq \frac{1}{4000}$. This simple fact is the bridge between the equilibrium concept and the semi-group action on X.

With respect to an ϵ-equilibrium for sufficiently small enough ϵ, there are two aspects of the game very important to our following arguments, First, where the strategies in an approximate equilibrium are not mixing, they tend to fall into the parity rule and stay there. Second, mixing is strongly discouraged by the structure of the payoffs. This dynamic is formalised in the next two lemmas.

Lemma 1: For every $x \in X$, regardless of the strategy of the G_0 player, either one or the other corresponding Player R_1 or R_2 at (r, x) has a preference of at least 80 to choose either $\mathbf{a_0}$ or $\mathbf{a_1}$ over the other action. Let $x, y \in X$ be any two points in X and let z_0 and z_1 be the twins where $T_1(z_i) = x$ and $T_2(z_i) = y$ for $i = 0, 1$ and $z_0^e = 0$ and $z_1^e = 1$. If one of R_1 or R_2 is not mixing at the corresponding (r, x) or (r, y), respectively, then Player G_0 at either (g, z_0) or at (g, z_1) has a preference of at least 80 to choose either $\mathbf{b_0}$ or $\mathbf{b_1}$ over the other action.

Proof: Without loss of generality assume that $x^e = 0$ and that the Player G_0 at (g, x) chooses $\mathbf{b_0}$ with probability at least $\frac{1}{2}$. By choosing $\mathbf{a_1}$ the R_1 player would get no more than 50 and by choosing $\mathbf{a_0}$ the R_1 player would get at least 150. On the other hand, if the Player G_0 at (g, x) chooses $\mathbf{b_1}$ with probability at least $\frac{1}{2}$ then the R_2 player would get no more than 50 by choosing $\mathbf{a_0}$ and at least 150 by choosing $\mathbf{a_1}$.

Next, due to symmetries, it suffices to consider the two cases of the R_2 player choosing $\mathbf{a_0}$ with probability no more than $\frac{1}{20}$ and the R_2 player choosing $\mathbf{a_1}$ with probability no more than $\frac{1}{20}$.

Let $w \leq \frac{1}{20}$ be the probability that the R_2 player chooses $\mathbf{a_0}$. We break this case into two sub cases, where Player R_1 chooses $\mathbf{a_0}$ with at least $\frac{3}{5}$ and where Player R_1 chooses $\mathbf{a_0}$ with at most $\frac{3}{5}$. If Player R_1 chooses $\mathbf{a_0}$ with at least $\frac{3}{5}$ then the G_0 player at (g, z_1) gets at least 570 for playing $\mathbf{b_0}$ and no more than $400(1 - w) + 2000w$ for playing $\mathbf{b_1}$, which reaches a maximum of 480 at $w = \frac{1}{20}$. If Player R_1 chooses $\mathbf{a_1}$ with at least $\frac{2}{5}$ then the G_0 player at (g, z_0) gets at least 760 from choosing $\mathbf{b_0}$ and no more than $600(1 - w) + 2000w$ for playing $\mathbf{b_1}$, which reaches a maximum of 670 at $w = \frac{1}{20}$.

Now let $w \leq \frac{1}{20}$ be the probability that the R_2 player chooses $\mathbf{a_1}$. We break this case into two sub cases, where Player R_1 chooses $\mathbf{a_0}$ with at least $\frac{3}{5}$ and where Player R_1 chooses $\mathbf{a_0}$ with at most $\frac{3}{5}$. If Player R_1 chooses $\mathbf{a_0}$ with at least $\frac{3}{5}$ then the G_0 player at (g, z_1) get at least 1140 from choosing $\mathbf{b_1}$ and no more than 820 by choosing $\mathbf{b_0}$. If Player R_1 chooses $\mathbf{a_1}$ with at least $\frac{2}{5}$ then the G_0 player at (g, z_0) gets at least 780 from choosing $\mathbf{b_1}$ and no more than $600(1-w)+2000w$ from choosing $\mathbf{b_0}$, which reaches a maximum of 670 at $w = \frac{1}{20}$. □

The consequence of Lemma 1 is that the players are hardly ever mixing in an approximate equilibrium. That is formalised in the next lemma.

Lemma 2: In any Borel measurable $\frac{1}{1000}$-equilibrium of the game, the G_0 player mixes with probability less than $\frac{16}{10{,}000}$ and the parity rule holds for all but at most $\frac{4}{1000}$ of the space X.

Proof: Let B_1 be the subset of X such that the R_1 player is mixing at the points labelled r and let B_2 be the subset of X such that the R_2 player at the points labelled r is mixing. Let $c = m(A_M)$, $a = m(B_1)$ and $b = m(B_2)$. As the $T_1 z$ and $T_2 z$ are distributed independently as variables of z, in an $\frac{1}{1000}$ equilibrium the following holds: $c \leq \frac{1}{4000} + ab + \frac{1}{2}\big(a+b\big)$, where the $\frac{1}{4000}$ refers to the maximum probability for the G_0 player to choose an action with at least $\frac{1}{20}$ probability that is suboptimal by a quantity of at least 80, the ab refers to the probability that both players R_i are mixing at both points $(r, T_i z)$ for $i = 1, 2$, and $\frac{a+b}{2}$ refers to the probability (from Lemma 1) that Player G_0's two actions give payoffs that are within 80 of each other for one but not both of the twins z_0 and z_1 (where one or the other of R_1 at $(r, T_1 z_j)$ or R_2 at $(r, T_2 z_j)$ are mixing, but not both). By Lemma 1, the probability of both R_1 mixing at (r, x) and R_2 mixing at (r, x) cannot exceed $\frac{1}{4000}$. From this we conclude that $a + b \leq c + \frac{1}{1000}$, considering also the possibility that G_0 is not mixing at (g, x) nevertheless one or the other of the R_i players at (r, x) is mixing.

From $ab \leq \frac{1}{4}(a+b)^2$, and the above, we get the quadratic $0 \leq c^2 - \frac{999}{500}c + \frac{3{,}001}{1{,}000{,}000}$. After completing the square we get that $|c - \frac{999}{1000}| \geq \sqrt{.995}$. Since c cannot be greater than 1 we are left with $c < .999 - .9974 = .0016$. The probability that the parity rule is not followed for a $z \in X$ is no more than the probability of the G_0 player mixing at either $(g, T_1 z)$ or $(g, T_2 z)$ plus the probability that the R_1 player at $(r, T_1 z)$, the R_2 player at $(r, T_2 z)$ or the G_0 player at (g, z) is not properly responding to the corresponding non-mixing behaviour. These probabilities sum to .00395. . □

Next we show it is impossible for there to exist a $\frac{1}{1000}$ equilibrium Borel measurable equilibrium, using the regularity of the measure.

Let $\mathcal{C}_n$ be the set of cylinders of depth n, where the two cylinders defined by the values $x^e = 0$ and $x^e = 1$ have depth 0. With $2^{n+1} - 1$ words of length n or less the cardinality of $\mathcal{C}_n$ is $2^{2^{n+1}-1}$ and $m(c) = 2^{-2^{n+1}+1}$ for all $c \in \mathcal{C}_n$. Recall the definition of the sets A_0 and A_1. For every $c \in \mathcal{C}_n$ and $i = 0, 1$ let $w_i(c)$ be the conditional probability $m(A_i \cap c)/m(c)$. For every cylinder c define $\eta(c) := \min_{i=0,1} w_i(c)$ and let $r(c)$ be the conditional probability in the cylinder c of belonging to the set where the parity rule does not hold.

In the next lemma, we show that the parity rule is a powerful force to equalise the probabilities for both actions $\mathbf{b_0}$ and $\mathbf{b_1}$. This cannot be guaranteed for all cylinders, due to the small yet persistent probability that the parity rule doesn't hold. But it does hold in general for most cylinders, regardless of the depth. Two free generators and the dual causation implicit in the parity rule force this equalisation.

Lemma 3: In any Borel measurable $\frac{1}{1000}$-equilibrium of the game, the average $q_i = \frac{\sum_{c \in \mathcal{C}_i} \eta(c)}{|\mathcal{C}_i|}$ is at least $\frac{1}{3}$ for every i.

Proof: The proof is by induction. There are two elements in $\mathcal{C}_0$ and eight elements in $\mathcal{C}_1$. Let c_0 and c_1 be the two elements of $\mathcal{C}_0$. We observe that

$$c_0 = \bigcup_{i,j=0,1} \{x \in X \mid T_1(x) \in c_i,\ T_2(x) \in c_j\} \cap \{x \in X \mid x^e = 0\} \text{ and}$$

$$c_1 = \bigcup_{i,j=0,1} \{x \in X \mid T_1(x) \in c_i,\ T_2(x) \in c_j\} \cap \{x \in X \mid x^e = 1\}.$$

Let x_0 and x_1 be two points such that $x_0^e = 0$, $x_1^e = 1$, and $x_0^U = x_1^U$ for every other $U \neq e$. However membership in A_0 or A_1 is determined by $T_1 x_i$ and $T_2 x_i$, the parity rule requires the opposite membership for x_j when $j \neq i$. As the parity rule must hold in a set of size at least $1 - \frac{1}{250}$, it follows that in the whole space the probability given to both A_0 and A_1 must be approximately the same. More precisely, as at least $\frac{124}{125}$ of the points of X are such that both it and its twin obey the parity rule, the probability must be at least $\frac{62}{125}$ for both A_0 and A_1. Now let c be either c_0 or c_1. As c is created by either the e position being 0 or 1 and the four combinations of c_0 and c_1 in both direction T_1 and T_2, whatever are the probabilities given for the two $w_i(c_0)$ and the two $w_i(c_1)$, the fact that $\frac{w_i(c_0)+w_i(c_1)}{2} \geq \frac{62}{125}$ for both $i = 0, 1$, implies that the conditional probability given to both A_0 and A_1 at c must be at least $2(\frac{62}{125})^2 - \frac{1}{125} \geq .48$.

We assume the claim is true for q_i. Every $t \in \mathcal{C}_{i+1}$ is created through the combination of a pair $c, d \in \mathcal{C}_i$ with a determination of 0 or 1 for the e position (though this determination will play no rule in the following argument). Let i_c be the label for the action $\mathbf{b_{i_c}}$ that is less frequent at c (defining the value $\eta(c)$), and define i_d the same way. Let j be the label for the action following from the parity rule determined by the e position and the combination of the i_c label with the dominant label $k \neq i_d$ at d or the i_d label with the dominant $l \neq i_c$ label at c (however that may be determined by the e position). If $r(c) = r(d) = r(t) = 0$ then the parity rule would give $\mathbf{b_j}$ exactly $\eta(c)(1-\eta(d)) + \eta(d)(1-\eta(c))$, as it would give the other action the greater quantity $(1-\eta(c))(1-\eta(d)) + \eta(c)\eta(d)$. Due to the influence of the quantities $r(c), r(d), r(t)$ we cannot say for sure that j is the action less taken at t. But we can say that $\eta(t) \geq -r(t) + \eta(c)(1-\eta(d)-r(d)) + \eta(d)(1-\eta(c)-r(c)) \geq \eta(c) + \eta(d) - 2\eta(c)\eta(d)) - r(t) - \frac{r(c)+r(d)}{2}$. But with $\sum_{c \in \mathcal{C}_j} r(c) \leq \frac{1}{250}|\mathcal{C}_j|$ for all j and the identical calculation conditioned for both halves $\{x \mid x^e = 0\}$ and $\{x \mid x^e = 1\}$ it follows that $q_{i+1} \geq -\frac{1}{125} + \frac{1}{|\mathcal{C}_i|^2} \sum_{c,d \in \mathcal{C}_i} \eta(c) + \eta(d) - 2\eta(c)\eta(d) = -\frac{1}{125} + \frac{1}{|\mathcal{C}_i|} \sum_{c \in \mathcal{C}_i} \eta(c) + \frac{1}{|\mathcal{C}_i|} \sum_{d \in \mathcal{C}_i} \eta(d) + \frac{1}{|\mathcal{C}_i|^2} (\sum_{c \in \mathcal{C}_i} \eta(c))(\sum_{d \in \mathcal{C}_i} \eta(d)) = -\frac{1}{125} + 2q_i - 2q_i^2$. By induction we conclude that $q_{i+1} \geq \frac{4}{9} - \frac{1}{125} > \frac{1}{3}$. □

Theorem 1: There can be no Borel measurable $\frac{1}{1000}$-equilibrium.

Proof: With $\eta(c)$ defined as in the proof of Lemma 3, for the mutually exclusive measurable sets A_0, A_1 of X it would follow from the regularity of the measure μ that $\lim_{n \to \infty} q_n = \lim_{n \to \infty} \frac{\sum_{c \in \mathcal{C}_n} \eta(c)}{|\mathcal{C}_n|} = 0$. But by Lemma 3 it never falls below $\frac{1}{3}$. □

As the probability distribution μ is regular, and the question concerns approximate equilibria, the negative result persists when we consider strategies that are measurable with respect to the completion of μ.

Notice that where the players are obeying the parity rule, even approximately so, the location where the payoff to Player G_0 is close to 2000 or to 1000 is determined by one or the other of the two other players, by Player R_1 in the half $\{x \in X \mid x^e = 0\}$ and by Player R_2 in the half $\{x \in X \mid x^e = 1\}$. Where the parity rule holds the measurability of the payoff of Player G_0 implies the measurability of the strategies of the R_i players and hence also the measurability of the strategy of the G_0 player in response. For the G_0

player to have an "equilibrium payoff" by some interpretation this payoff structure shows that one must define that concept quite remotely from the existence of Harsanyi ϵ-equilibria.

4 Bayesian equilibria:

Extend the parity rule so that it requires from a colouring function $c : X' \to \{0,1\}$ that $c(x) = c(T_1(x)) + c(T_2(x)) + x^e$ (modulo 2) for some subset $X' \subset X$ with $T_i(X') \subseteq X'$ for both $i = 1, 2$. In this section we will show that we can colour the space $X = \{0,1\}^{G^+}$ modulo a null set N using only two colours: 1 and 0 or red and blue, respectively, so that the parity rule is obeyed, and furthermore extend this to an equilibrium of the game played on the whole space Ω.

Recall the definition, stated in the section 2, of the twins determined by some $x, y \in X$. We say that a subset A of X is closed if for every pair x, y in A the twins determined by x and y are also in A. By the *closure* $\overline{A}$ of a set $A \subseteq X$ we mean the smallest closed set containing A. We say that $A \subseteq X$ is *pyramidic* if $x \in A$ implies that $U(x) \in A$ for all $U \in G^+$. The main example of pyramidic set is $G^+(x)$, where x is an arbitrary element of X.

Notice that whenever a set is pyramidic that its closure is also pyramidic, and a colouring of a pyramidic set that obeys the parity rule can be extended in the natural way to a colouring of its closure that obeys the parity rule. The latter may fail if the set A is not pyramidic, meaning that there is some $x \in A$ such that not all of $G^+(x)$ is in A. The reason for potential failure is that the colour of x before the closure operation may contradict that implied through the closure of the subset $A \cap (G^+(x) \setminus \{x\})$. But if all of $G^+(x)$ is in A and has been coloured consistently already according to the parity rule, this cannot happen. Of special concern are the twins determined by some x and y, or any pair z_1, z_2 that differ in at most finitely many positions. Consider the situation that z_1 is in a pyramidic set P but z_2 is not in P. Through the colouring of $G^+(z_1) \subseteq P$ and its closure, a colour for z_2 is determined. One must colour the closure of a pyramic set P before moving on to colour anything else; otherwise one may colour twins the same colour and the inconsistency becomes apparent only later.

Define now the set

$$N = \{x \in X : U(x) = V(x), \text{ for some distinct } U, V \in G^+\}.$$

This set is a null set with respect to the product measure m on X. Indeed, for any given two distinct words $U, V \in G^+$, the equality implies an agreement on infinitely many coordinates, and there are only countably many words $U \in G^+$. Let X_1 be $X \setminus N$. Notice that $T_i(X_1) \subseteq X_1$ for both $i = 1, 2$.

We are ready to prove the following lemma:

Lemma 4: Assuming the Axiom of Choice, there exists a colouring of $\overline{X_1}$ using the two colours $\{0, 1\}$ which satisfies the parity rule.

Proof: Let x_0 be any element of X_1 and obtain the set $P_0 := G^+(x_0)$. We define now a colouring of P_0 as follows:

(i) colour all the points $T_1U(x_0)$ in red, where $U \in G^+(x_0)$ and U is the identity or begins on the right with T_1;

(ii) colour all the points $T_2V(x_0)$ in blue, where $V \in G^+(x_0)$ is the identity or begins on the right with T_2;

(iii) colour the remaining points of the pyramid P_0 in the way that they satisfy the parity rule.

After colouring all the points of P_0, extend the colouring to all the points in the closure $\overline{P_0}$ of P_0.

Next create a partial ordering on colourings of pyramidic and closed subsets of X that obey the parity rule, with one colouring greater than another if the subset is larger and their colourings agree on their common intersection (the smaller subset). Any tower of such colourings will define a colouring that obeys the parity rule. As Zorn's Lemma implies that there is a maximal element, it suffices to show that maximality implies that all of X_1 has been coloured. Let P be any maximal pyramidic and closed subset of $\overline{X_1}$ with a colouring that obeys the parity rule. Assume that P does not include all of X_1. Let x be a member of X_1 that is not in P.

We say that x has a *hitting point in* P if $U(x) \in P$ for some $U \in G^+$ and whenever $U = VW$ and W is not the identity then $V(x) \notin P$.

Now we have the two cases:

Case 1) x has no hitting point with respect to P. Then we colour the closure $G^+(x)$ in the same way as the initial pyramidic set P_0.

Case 2) x has a hitting point in P. Then colour the elements of $G^+(x)$ taking into an account the colours of the hitting points. Notice that the closure of P implies that if Ux is not in P then one of UT_ix is also not in P, so that if UT_ix is a hitting point then UT_jx is not a hitting point, for $i \neq j$. This allows us to colour x arbitrarily and then move downward in a way consistent with the parity rule, with the colouring of UT_ix determined already only if, for $j \neq i$, the point UT_jx is a hitting point or UT_jx had just been coloured arbitrarily.

And then we colour the closure of $G^+(x) \cup P$ according to the parity rule for a larger set that is closed, pyramidic, and consistent with both the parity rule and the pre-existing colouring. □

Theorem 2: Assuming the Axiom of Choice, there exists a Bayesian equilibrium on all of Ω such that the strategies are pure (all weight on one action) for all but a subset of measure zero.

Proof: Following on from the proof of Lemma 4, we assume a colouring of $\overline{X_1}$ that obeys the parity rule. For all three players G_0, R_1 and R_2 define pure strategies on $D \times \overline{X_1}$ accordingly, with the action of the G_0 player at (g, x) and (r, x) determined by the colouring of x and the actions of the R_i player at (r, x) and $\{g\} \times T_i^{-1}(x)$ copying the colouring of x. With x any point in $X \backslash \overline{X_1}$, let Γ_x be the game defined on $D \times G^+(x) \backslash \overline{X_1}$ such that the strategies on $D \times (G^+(x) \cap \overline{X_1})$ are fixed by the above colouring. As the game has only countably many positions, by Simon (2003) there is a Nash equilibrium defined on the game Γ_x (from sequences of Nash equilibria defined on finite subsets). Notice that it defines an equilibrium when including those already fixed strategies on $D \times (G^+(x) \cap \overline{X_1})$ (as the players at these points in $D \times (G^+(x) \cap \overline{X_1})$ are not influenced by the behaviour done elsewhere). Extend this equilibrium to an equilibrium on the set $D \times \overline{(G^+(x) \cup \overline{X_1})}$ through optimal responses (noticing that nothing done at a point y has any influence on the payoffs of a player at points Uy for any $U \neq e$). Those optimal responses could involve in every case all weight given to a single action. As with the proof of Lemma 4, we define a partial ordering on pyramidic and closed subsets P of X with equilibria defined on $D \times P$ such that the partial ordering requires that the equilibrium behaviour has to agree on the common intersection. With P any closed and pyramidic subset for which an equilibrium is defined on $D \times P$, we define a game Γ_y^P on $D \times G^+(y) \backslash P$ for any $y \notin P$ and, following the same argument as above, show that it extends the equilibrium on $D \times P$ to one on $D \times \overline{G^+(y) \cup P}$. In this way, using Zorn's Lemma, we show that an equilibrium can be defined on all of Ω. □

There are some points in X for which any equilibrium requires a mixed strategy. Let x, y be the two points defined by $x^e = 0$, $y^e = 1$, $T_1(x) = y$, $T_2(x) = x$, $T_1(y) = x$ and $T_2(y) = x$. No matter how x is coloured, because $T_1(y) = T_2(y) = x$ and $y^e = 1$, y must be coloured with 1. But then $T_1(x) = y$, $T_2(x) = x$ and $x^e = 0$ forces x to be coloured differently from itself.

5 Conclusion: open questions

Is there an example of an ergodic game (Simon 2003) that has no Harsanyi ϵ-equilibrium for some positive ϵ? The examples of Simon (2003) and of Hellman (2014) were ergodic games, and ergodic games have Bayesian equilibria (Simon 2003). We believe the answer is yes and that it can be done through the action of a non-amenable group which defines the information structure of the players. With the example of this paper, there was a very strong mixing structure that kept the probability high for both actions at all cylinders. We believe that the weaker mixing structure from a group action would be sufficient to obtain the same result.

In the example of this paper, there are three players. Can the same result be accomplished with two players? Our initial belief was that the answer is yes, but now we are agnostic. With two players, by fixing the behaviour of one player the payoffs to both are affine in the behaviour of the other player. This is a restrictive condition, and yet we know of no theory preventing the sufficiently complex structure of causation.

Lastly, what is the relationship between Bayesian equilibria and the Banach-Tarski paradox? Let G be a group acting in a measure preserving way on a probability space X with measure μ and for every player i let G_i be a finite subgroup of G so that the information sets of Player i are the orbits of G_i and G is generated by the G_i. Is there a Bayesian game so defined such that for every Bayesian equilibrium and for every G-invariant finitely additive measure extending μ the equilibrium is not measurable with respect to the finitely additive measure? On this question we were initially agnostic, but now believe that it is possible.

6 References

Harsanyi, J. (1967) , Games with Incomplete Information Played by Bayesian Players, *Management Science*, Vol. 14, pp. 159-182.

Hellman, Z. (2014) , A Game with No Bayesian Approximate Equilibria, *Journal of Economic Theory*, Vol. 153, pp. 138-151.

Hellman, Z. and Levy, J. (2016) , Bayesian Games over Continuum Many States, *Theoretical Economics*, forthcoming.

Kechris, A., Solecki, S. Todorcevic, S. (1999) , Borel Chromatic Numbers, *Advances in Mathematics*, Vol 141, Issue 1, pp. 1-44.

Milgrom, P. and Weber, R. (1985) , Distribution Strategies for Games with Incomplete Information, *Mathematics of Operations Research*, Vol. 10, No. 4, pp. 619-632.

Simon, R. S. (2003) , Games of Incomplete Information, Ergodic Theory, and the Measurability of Equilibria, *Israel J. Math.*, Vol. 138, No. 1, pp. 73-92.