Jordan Learning Sequences and Proof of PS1 Equilibrium Proposition

Let G be a strategic form game of incomplete information, in which $I = \{1, \dots, I\}$ is the set of players. For each $i \in I$, let A_i denote the set of player i's available actions a_i; let K_i denote his set of possible types k_i; and let π_i denote his payoff function, with argument $(a, k_i) \in A \times K_i$, where $A = \times_{i \in I} A_i$. We also write π_i for i's expected payoff function. Each player knows his own type; let β_i denote the other players' (common) beliefs about player i's type, i.e., β_i is a probability measure on K_i. (We thus assume that the players believe the types to be distributed independently.)

Let G^* be an extensive form game defined by T repeated plays of G, with players observing, by the end of stage t, all actions taken at stage t. For each $i \in I$, let H_i denote the set of all player i's information sets h_i in the game G^*. There is no requirement that G^* be a game of perfect recall.

For each $t = 1, \dots, T$, let $\overline{a}^t = (a^1, a^2, \dots, a^t)$ denote a history of play through the first t periods of play; $\overline{a}^t$ specifies the action profile a^τ chosen at each stage τ. Let $\overline{a}^0$ denote the null history, the history prior to play at $t = 1$. For any $\overline{a}^t$, let $H_i(\overline{a}^t)$ denote the set of information sets player i could reach at his next play, following history $\overline{a}^t$, and let $H(\overline{a}^t) = \times_{i \in I} H_i(\overline{a}^t)$. Note that $H_i(\overline{a}^0) = K_i$, and that for every history $\overline{a}^t$, $H_i(\overline{a}^t)$ contains one element for every element of K_i, *i.e.*, it contains one information set for each of player i's types.

For each player $i \in I$, a behavior strategy for i is a function b_i from H_i, the set of all his possible information sets, into A_i, the set of his stage game actions. Let b be a profile of behavior stratgeies in G^*. For any player $i \in I$ and any $\overline{a}^t$, let $b_i[\overline{a}^t]$ denote the restriction of b_i to $H_i(\overline{a}^t)$, and let $b[\overline{a}^t] = (b_1[\overline{a}^t], \dots, b_I[\overline{a}^t])$, the corresponding profile of such restrictions.

Definition 1: A profile b of behavior strategies for G^* is a **Jordan profile** if it satisfies the following conditions:

(1) $b[\overline{a}^0]$ is a Bayesian equilibrium for the stage 1 game, the game defined by the profile π of payoff functions and the profile β^1 of initial beliefs; and

(2) For each $t = 1, \dots, T-1$ and for each history $\overline{a}^t$ such that $a^\tau \in$ support $b[\overline{a}^{\tau-1}]$ for

$\tau = 1, \ldots, t$, the profile $b[\overline{a}^t]$ is a Bayesian equilibrium of the game at stage $t+1$ defined by the profile π of payoff functions and the profile β^{t+1} of beliefs formed by Bayesian updating of β^t according to action profile a^t and behavior strategy profile $b[\overline{a}^{t-1}]$ — *i.e.*,

$$\forall k' \in K_i : \quad \beta_i^{t+1}(k') = \frac{\beta_i^t(k')b_i[\overline{a}^{t-1}](a_i^t, k')}{\sum_{k \in K_i} \beta_i^t(k)b_i[\overline{a}^{t-1}](a_i^t, k)}.$$

Note that in both (1) and (2), the profile of payoff functions is π, the components of which are the *stage* game payoff functions — *i.e.*, the players are assumed in this definition to behave myopically. In the definitions of perfect Bayesian equilibrium, sequential equilibrium, etc., conditions (1) and (2) appear, but the profile π is replaced with the profile of payoff functions for the *continuation* game defined at the profile $H(\overline{a}^t)$, *i.e.*, the continuation game at stage t.

Note too that the definition of a Jordan profile places no restrictions on behavior strategies at information sets that cannot be reached at stage t, given the profile's behavior strategies at stages prior to stage t. Thus, there are typically many Jordan profiles that are equivalent, in the sense that their behavior strategies are identical at all information sets they can reach. We refer to such an equivalence class as a Jordan Learning Sequence:

Definition 2: Two Jordan profiles are **equivalent** if they coincide at every information set which satisfies the conditions in Definition 1. An equivalence class of Jordan profiles is a **Jordan Learning Sequence**.

Thus, in order to describe a Jordan Learning Sequence it is sufficient to proceed recursively through the game G^* just as in the conditions in Definition 1, describing the behavior strategies at the information sets that can be reached — *i.e.*, describing the Bayesian equilibria (and the updated beliefs that support them) at each version of the stage game that can appear along the process's path.

We are now prepared to prove the PS1 Equilbrium Proposition (Section 7), which states that in the particular game G^* defined by Figure 5, with payoffs that are the sum of the payoffs at the two stage games, every PS1 equilibrium of G^* yields a Jordan Learning Sequence (*i.e.*, every PS1 equilibrium is a Jordan profile). There are four expected payoff functions that will be required in order to define equilibrium in the extensive form of G^*, one for each player in each of his two types. We write one of the four expected payoff functions in detail here, the Row player's expected payoff function when he is Type 0; the arguments are the Row and Column behavior strategies $r = (r_1, r_2, r_3, r_4, r_5, r_6)$ and $c = (c_1, c_2, c_3, c_4, c_5, c_6)$, as defined in Section 4.

$$\pi^{G^*}_{Row,0}(r, c) \;=\; \frac{1}{2}\pi^{G^*}_{Row,0}(r, c \mid g = 00) + \frac{1}{2}\pi^{G^*}_{Row,0}(r, c \mid g = 01),$$

where

$$\begin{aligned}\pi^{G^*}_{Row,0}(r, c \mid g = 00) \;=\;& \pi_{Row,0}(r_1, c_1 \mid g = 00) + \\ & (1 - r_1)(1 - c_1)\pi_{Row,0}(r_3, c_3 \mid g = 00)] + (1 - r_1)c_1\pi_{Row,0}(r_4, c_3 \mid g = 00)] + \\ & r_1(1 - c_1)\pi_{Row,0}(r_3, c_4 \mid g = 00)] + r_1c_1\pi_{Row,0}(r_4, c_4 \mid g = 00)],\end{aligned}$$

and

$$\begin{aligned}\pi^{G^*}_{Row,0}(r, c \mid g = 01) \;=\;& \pi_{Row,0}(r_1, c_2 \mid g = 01) + \\ & (1 - r_1)(1 - c_2)\pi_{Row,0}(r_3, c_5 \mid g = 01)] + (1 - r_1)c_2\pi_{Row,0}(r_4, c_5 \mid g = 01)] + \\ & r_1(1 - c_2)\pi_{Row,0}(r_3, c_6 \mid g = 01)] + r_1c_2\pi_{Row,0}(r_4, c_6 \mid g = 01)].\end{aligned}$$

These expressions will be simplified considerably in the proof, because in a PS1 equilibrium we have $r_1 = 1$, $r_2 = 0$, and $c_1 = c_2 = 0$, so that most of the terms will vanish. The same is true for the other three expected payoff functions as well.

Proof of the JLS Equilbrium Proposition:

Let $(\mathring{r}, \mathring{c})$ be a PS1 profile in the game G^* — *i.e.*, $(\mathring{r}_1, \mathring{r}_2; \mathring{c}_1, \mathring{c}_2)$ is a pure strategy Bayesian equilibrium of the Period 1 game defined by the beliefs $(\beta^1_{Row}, \beta^1_{Col}) = ((\tfrac{1}{2}, \tfrac{1}{2}), (\tfrac{1}{2}, \tfrac{1}{2}))$. We have already seen that there is only one such BE: $(\mathring{r}_1, \mathring{r}_2; \mathring{c}_1, \mathring{c}_2) = (B, T; L, L) = (1, 0; 0, 0)$. Since that Period 1 Bayesian equilibrium admits two possible action profiles at $t = 1$, *viz.*, (B, L) and (T, L), it follows that there are two possible games of incomplete information at $t = 2$; we will refer to them as the (B, L)-game and the (T, L)-game. The updated beliefs for the (B, L)-game are $(\beta^2_{Row}, \beta^2_{Col}) = ((1, 0), (\frac{1}{2}, \frac{1}{2})$ — *i.e.*, Column now knows that Row is Type 0, while Row ascribes probability 1/2 to each of Column's types. The updated beliefs for the (T, L)-game are $(\beta^2_{Row}, \beta^2_{Col}) = ((0, 1), (\frac{1}{2}, \frac{1}{2}))$ — in this case Column knows that Row is Type 1.

We must show that if $(\mathring{r}, \mathring{c})$ is a PS1 equilibrium of G^*, then it is a Jordan profile — i.e., that the profile $(\mathring{r}_3, \mathring{r}_4, \mathring{r}_5, \mathring{r}_6; \mathring{c}_3, \mathring{c}_4, \mathring{c}_5, \mathring{c}_6)$ yields Bayesian equilibria in both the (B, L)-game and the (T, L)-game at $t = 2$. We proceed indirectly, assuming that $(\mathring{r}, \mathring{c})$ is *not* a Jordan profile, and showing that this implies that it is not an equilibrium of G^*. Assume then that $(\mathring{r}, \mathring{c})$ is not a Jordan profile, which (since we have already assumed that it *is* a PS1 profile) means that it fails to yield a BE in at least one of the games at $t = 2$, the (B, L)-game and/or the (T, L)-game. Thus, in at least one of those two games at least one of the two players has a profitable deviation from the restriction of $\mathring{r}$ or $\mathring{c}$ to the information sets in that game. We consider each of these four cases in turn, determining a set of six inequalities, at least one of which must be satisfied if $(\mathring{r}, \mathring{c})$ fails to yield a Bayesian equilibrium in either the (B, L)-game or the (T, L)-game.

In the (B, L)-game, assuming that Column is playing $\mathring{c}$, Row's expected payoff function is

$$\frac{1}{2}\pi_{Row,0}(r_3, \mathring{c}_4 |\ g = 00) + \frac{1}{2}\pi_{Row,0}(r_3, \mathring{c}_6 |\ g = 01).$$

Consequently, a profitable deviation for Row in the (B, L)-game would be an r_3 for which

$$\frac{1}{2}\pi_{Row,0}(r_3, \mathring{c}_4 |\ g = 00) + \frac{1}{2}\pi_{Row,0}(r_3, \mathring{c}_6 |\ g = 01) > \frac{1}{2}\pi_{Row,0}(\mathring{r}_3, \mathring{c}_4 |\ g = 00) + \frac{1}{2}\pi_{Row,0}(\mathring{r}_3, \mathring{c}_6 |\ g = 01). \tag{1}$$

Similarly, in the (T, L)-game a profitable deviation for Row would be an r_5 for which

$$\frac{1}{2}\pi_{Row,1}(r_5, \mathring{c}_3 |\ g = 10) + \frac{1}{2}\pi_{Row,1}(r_5, \mathring{c}_5 |\ g = 11) > \frac{1}{2}\pi_{Row,1}(\mathring{r}_5, \mathring{c}_3 |\ g = 10) + \frac{1}{2}\pi_{Row,1}(\mathring{r}_5, \mathring{c}_5 |\ g = 11). \tag{2}$$

For the Column player, there are two information sets to consider in each game, producing altogether four inequalities. In the (B, L)-game, a profitable deviation for Column would be either a c_4 for which

$$\pi_{Col,0}(\overset{\circ}{r}_3, c_4 \mid g = 00) > \pi_{Col,0}(\overset{\circ}{r}_3, \overset{\circ}{c}_4 | \ g = 00), \tag{3}$$

or else a c_6 for which

$$\pi_{Col,1}(\overset{\circ}{r}_3, c_6 \mid g = 01) > \pi_{Col,1}(\overset{\circ}{r}_3, \overset{\circ}{c}_6 | \ g = 01). \tag{4}$$

Similarly, in the (T, L)-game a profitable deviation for the Column player would be either a c_3 for which

$$\pi_{Col,0}(\overset{\circ}{r}_5, c_3 \mid g = 10) > \pi_{Col,0}(\overset{\circ}{r}_5, \overset{\circ}{c}_3 | \ g = 10), \tag{5}$$

or else a c_5 for which

$$\pi_{Col,1}(\overset{\circ}{r}_5, c_5 \mid g = 11) > \pi_{Col,1}(\overset{\circ}{r}_5, \overset{\circ}{c}_5 | \ g = 11). \tag{6}$$

We will show that if any of the six inequalities above holds (*i.e.*, if $(\overset{\circ}{r}, \overset{\circ}{c})$ does not yield a Bayesian equilibrium in each of the two games at $t = 2$), then $(\overset{\circ}{r}, \overset{\circ}{c})$ is not an equilibrium of the game G^*. Note that in order for $(\overset{\circ}{r}, \overset{\circ}{c})$ to be an equilibrium of the game G^*, a deviation from $\overset{\circ}{r}$ to any other Row behavior strategy r in G^* cannot increase Row's expected payoff $\pi^{G^*}_{Row,0}$ when he is Type 0, *i.e.*,

$$\pi^{G^*}_{Row,0}(\overset{\circ}{r}, \overset{\circ}{c}) - \pi^{G^*}_{Row,0}(r, \overset{\circ}{c}) \quad \geq \quad 0,$$

nor can such a deviation from $\overset{\circ}{r}$ increase his expected payoff $\pi^{G^*}_{Row,1}$ when he is Type 1. Similarly, a deviation from $\overset{\circ}{c}$ to another behavior strategy c cannot increase either $\pi^{G^*}_{Col,0}$ or $\pi^{G^*}_{Col,1}$ if $(\overset{\circ}{r}, \overset{\circ}{c})$ is to be an equilibrium of G^*. We will show, however, that if any of the six inequalities (1) - (6) is satisfied, then there *is* a deviation in G^* for one of the players that will increase either his expected payoff when he is Type 0 or his expected payoff when he is Type 1.

The first of these four expected payoff functions, $\pi^{G^*}_{Row,0}$, is as follows, when Column is playing strategy $\overset{\circ}{c}$ (*cf.* the complete payoff function, which appears just before the beginning of the proof):

$$\begin{aligned}
\pi^{G^*}_{Row,0}(r,\overset{\circ}{c}) &= \frac{1}{2}\pi^{G^*}_{Row,0}(r,\overset{\circ}{c}|\ g=00)+\frac{1}{2}\pi^{G^*}_{Row,0}(r,\overset{\circ}{c}|\ g=01)\\
&= \frac{1}{2}[\pi_{Row,0}(r_1,\overset{\circ}{c}_1|\ g=00)+\pi_{Row,0}(r_3,\overset{\circ}{c}_4|\ g=00)]+\\
&\quad \frac{1}{2}[\pi_{Row,0}(r_1,\overset{\circ}{c}_2|\ g=01)+\pi_{Row,0}(r_3,\overset{\circ}{c}_6|\ g=01)]\\
&= \frac{1}{2}[\pi_{Row,0}(r_1,\overset{\circ}{c}_1|\ g=00)+\pi_{Row,0}(r_1,\overset{\circ}{c}_2|\ g=01)]+\\
&\quad \frac{1}{2}[\pi_{Row,0}(r_3,\overset{\circ}{c}_4|\ g=01)+\pi_{Row,0}(r_3,\overset{\circ}{c}_6|\ g=01)]
\end{aligned}$$

Now suppose there is a strategy r_3 at Row's third information set that satisfies inequality 1. Then substituting the behavior strategy $r' = (\overset{\circ}{r}_1,\overset{\circ}{r}_2,r_3,\overset{\circ}{r}_4,\overset{\circ}{r}_5,\overset{\circ}{r}_6)$ into the expression above for $\pi^{G^*}_{Row,0}(r,\overset{\circ}{c})$ clearly yields an expected payoff greater than $\pi^{G^*}_{Row,0}(\overset{\circ}{r},\overset{\circ}{c})$: the only term that is changed is the one involving r_3, which is increased by the deviation. Thus, if a strategy r_3 satisfies 1, the Row player has a profitable deviation in G^* from $(\overset{\circ}{r},\overset{\circ}{c})$, and thus $(\overset{\circ}{r},\overset{\circ}{c})$ is *not* an equilibrium of the game G^*.

Similarly, if there is a strategy r_5 at Row's fifth information set that satisfies inequality 2, then substituting the behavior strategy $r' = (\overset{\circ}{r}_1,\overset{\circ}{r}_2,\overset{\circ}{r}_3,\overset{\circ}{r}_4,r_5,\overset{\circ}{r}_6)$ into the expression for $\pi^{G^*}_{Row,1}(r,\overset{\circ}{c})$ yields an expected payoff greater than $\pi^{G^*}_{Row,1}(\overset{\circ}{r},\overset{\circ}{c})$: the only term that is changed is the one involving r_5. Thus, if a strategy r_5 satisfies 2, $(\overset{\circ}{r},\overset{\circ}{c})$ is not an equilibrium of the game G^*.

The same argument applies to potential deviations by the Column player. If there is a strategy c_4 that satisfies inequality 3, then a behavior strategy c' in G^* which changes just the fourth component of $\overset{\circ}{c}$ to c_4 will increase just the term involving c_4 in Column's expected payoff $\pi^{G^*}_{Col,0}$, so that c' will be a profitable deviation from $\overset{\circ}{c}$. And if there is a strategy c_6 that satisfies inequality 4, then a behavior strategy c' in G^* which changes just the sixth component of $\overset{\circ}{c}$ to c_6 will increase just the term involving c_6 in $\pi^{G^*}_{Col,0}$, so that this c' too will be a profitable deviation from $\overset{\circ}{c}$. Strategies c_3 or c_5 that satisfy 5 or 6 can similarly be used to construct a profitable deviation c' that increases just one term in the expression for $\pi^{G^*}_{Col,1}$. Thus, in each of these four cases for the Column player, if there is a deviation that upsets the proposed Bayesian equilibrium of either the (B, L)-game or the (T, L)-game, then $(\overset{\circ}{r}, \overset{\circ}{c})$ is not an equilibrium of G^*.

Thus, we have established that if $(\overset{\circ}{r}, \overset{\circ}{c})$ is a PS1 profile of G^* but not a Jordan profile ($i.e.$, if in either the (B, L)-game or the (T, L)-game it fails to yield a Bayesian equilibrium), then $(\overset{\circ}{r}, \overset{\circ}{c})$ is not an equilibribium of G^*. Therefore, every PS1 equilibrium of G^* is a Jordan profile. $\square$