Notes on Sequential and Self Confirming Equilibrium

David Rahman, UCLA[*,†]

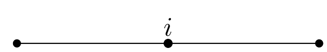
Winter, 2003

1 The Extensive Form of a Game

First I formally define a game in extensive form, and then I look at mixed and behavior strategies.

1.1 Formal Description of a Game

Perhaps obsessively, I will avoid reference to any objects that aren't derived explicitly from terminal nodes or players. I won't distinguish games that differ by redundant moves, i.e., branches like this.



Formally, we begin with a finite set Z of terminal nodes. Let $X \subset 2^Z$ be a family of subsets of Z. Call X a *rooted tree over* Z if it satisfies the following three conditions:

1. $Z \in X$ (there is a root node),
2. $\{z\} \in X$ for every $z \in Z$ (exhaustion of alternatives),
3. For any $x, y \in X$, if $x \cap y \neq \emptyset$ then either $x \subset y$ or $y \subset x$ (uniqueness of paths).[1]

We say that y is a *successor* of x if $y \subset x$. The interpretation of a node x is that it fully characterizes (and modulo redundant nodes is also characterized by) the set of possible terminal nodes that are reachable from x. This motivates defining Z as the root node, and $\{z\}$ as a terminal node.

*Please send any comments to dmr@ucla.edu.

[1] My convention is that $\subset$ denotes *weak* inclusion (so $x \subset x$).

For any two nodes $x \supset y$, the partial path from x to y is denoted and defined as

$$[x, y] := \{w \in X : x \supset w \supset y\}\,.$$

By definition, there is only one partial path from x to y. In particular, every terminal node $\{z\}$ has a unique path to it from the root node Z. The set of nonterminal nodes is $X \setminus Z$.[2]

Let $S(x) = \{y : [x, y] = \{x, y\}\}$ be the set of immediate successors of x.[3] A nodal decision is a partial path $[x, y]$ with $y \in S(x)$. A nodal decision problem is the set of possible decisions $[x, S(x)] = \{[x, y] : y \in S(x)\}$. Hence, each node is uniquely associated with a decision problem.

A *decision structure* for Z is a triple (X, H, A) satisfying:

1. X is a rooted tree over Z,
2. H is a partition of the nonterminal nodes $X \setminus Z$ such that for every $h \in H$ and $x, y \in h$, $|S(x)| = |S(y)|$, that is, x and y have the same number of immediate successors,
3. A is a correspondence mapping each information set $h \in H$ to a subset

$$A(h) \subset \prod_{x \in h} [x, S(x)]$$

   such that for any $x \in h$ and $y \in S(x)$, there is a unique $a \in A(h)$ with $a(x) = [x, y]$.

We now define players' tastes and their turns in the game. Let I be a set of players, with i's utility function $u_i : Z \to \mathbb{R}$ defined on terminal nodes. It remains is to allocate information sets to players.

An *allocation of turns and tastes to* I is a triple (u, T, μ) such that $u : I \times Z \to \mathbb{R}$ is a profile of utility functions, $T : H \to I \cup \{N\}$ is an allocation of *turns*[4] (i.e., information sets) to players, and μ belongs to

$$\prod_{h \in T^{-1}(N)} \Delta(A(h)),$$

the product space of probability measures over each of nature's possible moves. Let's denote by $H_i = T^{-1}(i)$ the collection of information sets that belong to player i.

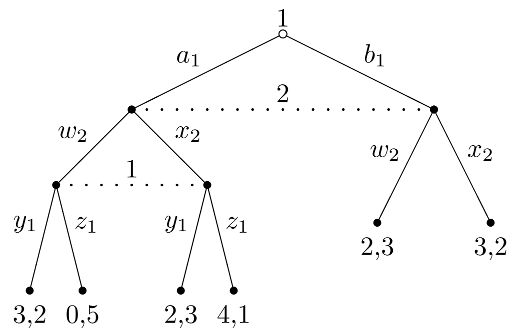
Definition 1.1 *An* extensive form game *is a tuple* $\Gamma = (Z, (X, H, A), I, (u, T, \mu))$ *such that*

1. (X, H, A) *is a decision structure for* Z*,*
2. (u, T, μ) *is an allocation of turns and tastes to* I*.*

Without further ado, let's consider a simple game and translate it into this formal framework.

[2] This is not $X \setminus \{Z\}$, which is the set of noninitial nodes.
[3] By condition 2, $S(x)$ is empty if and only if x is a terminal node.
[4] Player N is nature.

Example 1.2 *Consider the following extensive-form game.*



The game in Example 1.2 is described as follows. The set of terminal nodes is

$$Z = \{b_1w_2\ ,\ b_1x_2\ ,\ a_1w_2y_1\ ,\ a_1w_2z_1\ ,\ a_1x_2y_1\ ,\ a_1x_2z_1\}\,.$$

The rooted tree $X \subset 2^Z$ is constructed so that each node is represented by the possible terminal nodes from it. Hence,

$$\begin{aligned} X &= \{Z\ ,\ \{a_1w_2y_1\ ,\ a_1w_2z_1\ ,\ a_1x_2y_1\ ,\ a_1x_2z_1\}\ ,\ \{b_1w_2\ ,\ b_1x_2\}\ ,\\ &\quad \{a_1w_2y_1\ ,\ a_1w_2z_1\}\ ,\ \{a_1x_2y_1\ ,\ a_1x_2z_1\}\\ &\quad \{b_1w_2\}\ ,\ \{b_1x_2\}\ ,\ \{a_1w_2y_1\}\ ,\ \{a_1w_2z_1\}\ ,\ \{a_1x_2y_1\}\ ,\ \{a_1x_2z_1\}\}. \end{aligned}$$

Players 1 and 2 play this game. Since we can identify nodes on the game tree with partial paths on the tree from the root, there is no loss in *denoting* the set X of nodes by

$$X = \{\aleph\ ,\ a_1\ ,\ b_1\ ,\ a_1w_2\ ,\ a_1x_2\ ,\ b_1w_2\ ,\ b_1x_2\ ,\ a_1w_2y_1\ ,\ a_1w_2z_1\ ,\ a_1x_2y_1\ ,\ a_1x_2z_1\}\,,$$

where $\aleph$ stands for the root node. There are no chance nodes in this game. The information sets belonging players 1 and 2 are

$$\begin{aligned} H_1 &= \{\aleph, \{a_1w_2\ ,\ a_1x_2\}\}\,,\\ H_2 &= \{\{a_1\ ,\ b_1\}\}\,. \end{aligned}$$

For each player i, an element of H_i is an information set for i. The action correspondence A is given by the sets of partial paths below:[5]

$$\begin{aligned} &A(\aleph) = \{a_1,\ b_1\}\,,\\ &A(\{a_1\ ,\ b_1\}) = \{w_2\ ,\ x_2\}\,,\\ &A(\{a_1w_2\ ,\ a_1x_2\}) = \{y_1\ ,\ z_1\}\,. \end{aligned}$$

[5]Formally, $w_2 = \{a_1w_2\ ,\ b_1w_2\}$, etc.

1.2 Mixed and Behavior Strategies

There are two basic ways to think of strategic behavior in extensive form games. We will now define a mixed strategy and a behavior strategy.

Definition 1.3 *Fix a player $i \in N$. A* mixed strategy *for player i is any σ_i in*

$$\Sigma_i := \Delta\left(\prod_{h_i \in H_i} A(h_i)\right). \qquad \text{(Normal-form representation.)}$$

A behavior strategy *for player i is any π_i in*

$$\Pi_i := \prod_{h_i \in H_i} \Delta\left(A(h_i)\right). \qquad \text{(Multi-agent representation.)}$$

The set of pure strategies is the product space

$$C_i := \prod_{h_i \in H_i} A(h_i),$$

where each $c_i \in C_i$ is to be interpreted as a complete contingent plan *for player i.*

I will also write Σ_{-i}, Π_{-i}, Σ, Π to mean the obvious objects. (The product spaces across all players but i and across all players, respectively.) Let's find C_1 and C_2 in the example.

$$\begin{aligned} C_1 &= \{a_1\ ,\ b_1\} \times \{y_1\ ,\ z_1\} = \{(a_1, y_1), (a_1, z_1), (b_1, y_1), (b_1, z_1)\}, \\ C_2 &= \{w_2,\ x_2\}. \end{aligned}$$

A mixed strategy for player i is an element of $\Delta(C_i)$. Clearly, the mixed and behavior strategies for Player 2 coincide. An example of a mixed strategy for Player 1 is $\frac{1}{2}[(a_1, y_1)] + \frac{1}{2}[(b_1, z_1)]$.[6] A behavior strategy for player 1 is an element of

$$\Delta(\{a_1\ ,\ b_1\}) \times \Delta(\{y_1\ ,\ z_1\}).$$

An example of a behavior strategy is $(\frac{1}{2}[a_1] + \frac{1}{2}[b_1], \frac{1}{2}[y_1] + \frac{1}{2}[z_1])$. This strategy is clearly very different from the mixed strategy above. In the mixed strategy, Player 1 doesn't mix at all between y_1 and z_1.

Given a behavior strategy π_i, its *mixed-strategy representation* is the mixed strategy σ_i given by

$$\sigma_i(c_i) = \prod_{h_i \in H_i} \pi_i(c_i(h_i)).$$

We construct the mixed-strategy representation by assuming that behavior strategies involve independent mixing across information sets. For instance, the mixed-strategy representation of the behavior strategy $(\frac{1}{2}[a_1]+\frac{1}{2}[b_1], \frac{1}{2}[y_1]+\frac{1}{2}[z_1])$ is given by $\frac{1}{4}[(a_1, y_1)]+\frac{1}{4}[(a_1, z_1)]+\frac{1}{4}[(b_1, y_1)]+\frac{1}{4}[(b_1, z_1)]$.

[6] For any choice x, the notation $[x]$ stand for the lottery that yields x with unit probability.

Given a mixed-strategy *profile* σ and a terminal node $z \in Z$, the probability of z being reached under σ is denoted by $P(z|\sigma)$. Given an information state h_i, the probability of h_i being reached under σ is denoted by $P(h_i|\sigma)$.[7] Given a behavior strategy π_i, I will denote by $\hat{\sigma}_i(\cdot|\pi_i)$ the mixed representation of π_i. I will also write, for a behavior-strategy profile π, $P(x|\pi)$ to represent the probability that x is reached under π. Clearly, $P(x|\pi) = P(x|\hat{\sigma}(\cdot|\pi))$. In Example 1.2, let σ be the pure-strategy profile $((b_1, y_1), w_2)$. It is easy to verify that $P(z|\sigma) = 1$ if $z = b_1w_2$ and 0 otherwise. Similarly, $P(\{a_1\ ,\ b_1\}\,|\sigma) = 1$, but $P(\{a_1w_2\ ,\ a_1x_2\}\,|\sigma) = 0$.

Definition 1.4 *For each information set $h_i \in H_i$, and action $a_i \in A(h_i)$, let*

$$\begin{aligned} C_i^*(h_i) &= \{c_i \in C_i : \exists c_{-i} \text{ s.t. } P(h_i|c) > 0\}\,, \\ C_i^{**}(h_i, a_i) &= \{c_i \in C_i^*(h_i) : c_i(h_i) = a_i\}\,. \end{aligned}$$

For example, I compute the following sets.

$$\begin{aligned} C_1^* = C_1^*(\{a_1w_2\ ,\ a_1x_2\}) &= \{(a_1, y_1), (a_1, z_1)\}, \\ C_1^{**} = C_1^{**}(\{a_1w_2\ ,\ a_1x_2\}\,, y_1) &= \{(a_1, y_1)\}. \end{aligned}$$

Given a mixed strategy σ_i, a behavior strategy π_i is called a *behavioral representation* of σ_i if for every $h_i \in H_i$ and every $a_i \in A(h_i)$, the following equality holds:

$$\pi_i(h_i)(a_i)\left[\sum_{c_i \in C_i^*(h_i)} \sigma_i(c_i)\right] = \sum_{c_i \in C_i^{**}(h_i, a_i)} \sigma_i(c_i). \qquad (*)$$

Clearly, if the summation on the left is zero then the summation on the right is zero, too. Notice that if we read the terms of $(*)$ as

$$P(a_i|h_i) := \pi_i(h_i)(a_i), \qquad P(h_i|\sigma_i) := \sum_{c_i \in C_i^*(h_i)} \sigma_i(c_i), \qquad P(h_i, a_i|\sigma_i) := \sum_{c_i \in C_i^{**}(h_i, a_i)} \sigma_i(c_i),$$

then $(*)$ is just Bayes' rule: $P(a_i|h_i)P(h_i|\sigma_i) = P(h_i, a_i|\sigma_i)$. Hence, a behavioral representation of σ_i is any behavior strategy π_i that satisfies, given $h_i \in H_i$ and $a_i \in A(h_i)$,

$$\pi_i(h_i)(a_i) = \begin{cases} \frac{P(h_i, a_i|\sigma_i)}{P(h_i|\sigma_i)} & \text{if } P(h_i|\sigma_i) > 0, \\ \text{anything} & \text{otherwise,} \end{cases}$$

with the proviso that "anything" must add up to one on every h_i.[8] For example, consider the mixed strategies $\frac{1}{2}[(a_1, y_1)] + \frac{1}{2}[(b_1, z_1)]$ and $\frac{1}{2}[(a_1, y_1)] + \frac{1}{2}[(b_1, y_1)]$. They both have the same behavioral representation $(\frac{1}{2}[a_1] + \frac{1}{2}[b_1], [y_1])$, which in turn has the mixed-strategy representation $\frac{1}{2}[(a_1, y_1)] + \frac{1}{2}[(b_1, y_1)]$.

[7] Without chance nodes, if c is a pure-strategy profile, then $P(h_i|c)$ is either 0 or 1.

[8] I.e., behavioral representations are only defined up to zero-probability events, like conditional expectation.

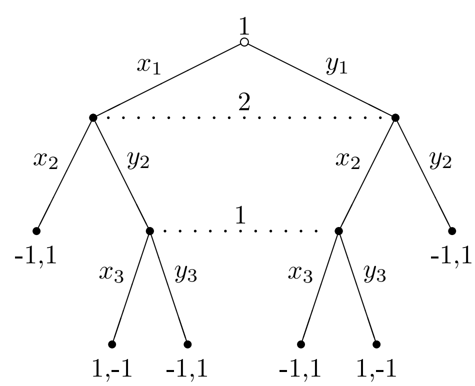
Definition 1.5 *Two mixed strategies $\sigma_i, \tau_i \in \Delta(C_i)$ are called* behaviorally equivalent *if they have the same behavioral representation whenever $P(h_i|\sigma_i)P(h_i|\tau_i) > 0$. They are called* payoff equivalent *if for every player $j \in N$ and any profile of strategies by i's opponents σ_{-i},*

$$u_j(\sigma_{-i}, \sigma_i) = u_j(\sigma_{-i}, \tau_i),$$

where u_j is Player j's utility function in the normal-form representation of the game.

To illustrate, consider the game in Example 1.6 below. This game doesn't have perfect recall: player 1 forgets his move in the first node. Consider the mixed strategies $\sigma_1 = \frac{1}{2}[(x_1, x_3)] + \frac{1}{2}[(y_1, y_3)]$ and $\tau_1 = \frac{1}{2}[(x_1, y_3)] + \frac{1}{2}[(y_1, x_3)]$. It is clear that they are behaviorally equivalent, with behavioral representation $(\frac{1}{2}[x_1] + \frac{1}{2}[y_1], \frac{1}{2}[x_3] + \frac{1}{2}[y_3])$. However, they're *not* payoff equivalent. For instance, given $[x_2]$, $u_1(\sigma_1, [x_2]) = 0$ yet $u_1(\tau_1, [x_2]) = -1$.

Example 1.6 *A game with imperfect recall.*



We will generally restrict ourselves to games with perfect recall, for the following reason.

Theorem 1.7 (Kuhn, 1953) *In every game with perfect recall, any two mixed strategies that are behaviorally equivalent are also payoff equivalent.*

See Myerson's book for a proof.

2 Equilibrium in Extensive Form Games

This section follows pretty closely Chapter 4 of Myerson's book.

2.1 Drawbacks in the Normal Form and Multi Agent Representations

In Section 1.2 we defined the normal-form (NFR) and multi-agent (MAR) representations of a game. Restricted to Nash equilibrium, one is left with the dilemma of too many equilibria in the NFR and unintuitive ones in the MAR. To illustrate, consider the game in Example 1.2.

	w_2	x_2
a_1y_1	$\mathbf{3}, 2$	$2, \mathbf{3}$
a_1z_1	$0, \mathbf{5}$	$\mathbf{4}, 1$
b_1y_1	$2, \mathbf{3}$	$3, 2$
b_1z_1	$2, \mathbf{3}$	$3, 2$

In this NFR, there are no pure Nash equilibria. However, there are many mixed Nash equilibria. For example, $\left(\frac{1}{2}[a_1y_1] + \frac{1}{2}(p[b_1y_1] + (1-p)[b_1z_1]), \frac{1}{2}[w_2] + \frac{1}{2}[x_2]\right)$ is a mixed NE if $0 \leq p \leq 1$. Trivially, there are too many (behaviorally equivalent) equilibria that lead to the same behavior strategy: $((\frac{1}{2}[a_1] + \frac{1}{2}[b_1], [y_1]), \frac{1}{2}[w_2] + \frac{1}{2}[x_2])$. On the other hand, MAR leads to a slightly different problem with Nash equilibrium. Working our way back from the end, let player 1 be the player choosing between y_1 and z_1, player 2 be player 2, and player 3 be the player at the root node.

	w_2	x_2
y_1	$\mathbf{3}, 2, \mathbf{3}$	$2, \mathbf{3}, \mathbf{2}$
z_1	$0, \mathbf{5}, 0$	$\mathbf{4}, 1, \mathbf{4}$

a_1

	w_2	x_2
y_1	$\mathbf{2}, \mathbf{3}, 2$	$\mathbf{2}, \mathbf{3}, \mathbf{2}$
z_1	$\mathbf{2}, 3, \mathbf{3}$	$\mathbf{2}, \mathbf{3}, 2$

b_1

The profile (y_1, x_2, b_1) is a pure NE. However, it seems to reflect a miscoordination between the two player 1's at different information sets. In this sense, the NE is unintuitive. This motivates the following definition.

Definition 2.1 *Let Γ be an extensive form game, with NFR Γ_N and MAR Γ_M. A behavior strategy profile π is called an* equilibrium of Γ *if π is a Nash equilibrium of Γ_M and the mixed representation $\hat{\sigma}$ of π is a Nash equilibrium of Γ_N.*

There is only one equilibrium in Example 1.2: $((\frac{1}{2}[a_1] + \frac{1}{2}[b_1], [y_1]), \frac{1}{2}[w_2] + \frac{1}{2}[x_2])$.

Proposition 2.2 *In games Γ with perfect recall, if σ is a NE of Γ_N then any behavior representation $\hat{\pi}$ of σ is a NE of Γ_M.*

Therefore, in games with perfect recall, an equilibrium always exists.[9]

[9] The perfect recall condition is *necessary*: in Example 1.6, Γ_N has only one NE, $(\frac{1}{2}[x_1x_3] + \frac{1}{2}[y_1y_3], \frac{1}{2}[x_2] + \frac{1}{2}[y_2])$. However, there is no behavioral representation of this profile that makes it a NE of Γ_M.

2.2 Sequential Rationality

Informally, a behavior strategy profile π is *sequentially rational for player i at history h_i* if $\pi_i(h_i) \in \Delta(A(h_i))$ is optimal amongst all alternatives that only change $\pi_i(h_i)$. Formally, we require that

$$\pi_i(h_i) \in \arg\max_{d_i(h_i)} \left\{ u_i([\pi_i : d_i(h_i)], \pi_{-i}|h_i) : d_i(h_i) \in \Delta(A(h_i)) \right\},$$

where $[\pi_i : d_i(h_i)]$ is the strategy that only replaces $\pi_i(h_i)$ with $d_i(h_i)$, and $([\pi_i : d_i(h_i)], \pi_{-i}|h_i)$ is the profile of conditional probabilities (over terminal nodes) that result from applying Bayes' rule to $([\pi_i : d_i(h_i)], \pi_{-i})$ conditional on h_i.

For now, we assume that h_i has *positive probability under* π. The caveat of positive probability constitutes a serious drawback. While in (statistical) decision theory zero-probability events are usually unimportant and ignored, in game theory zero-probability events are crucially important. We will address this issue shortly.

From the profile π, we may infer, for each terminal node z, the probability that z is reached under π, $P(z|\pi)$. For any nonterminal node history h_i and nonterminal node $x \in h_i$, since $P(A \cup B) = P(A) + P(B)$ when $A \cap B = \emptyset$, it follows that

$$P(x|\pi) = \sum_{z \in x} P(z|\pi), \qquad P(h_i|\pi) = \sum_{x \in h_i} P(x|\pi).$$

Therefore, we may induce from π player i's "beliefs" regarding the probability of being a node $x \in h_i$ conditional on reaching h_i:

$$P_\pi(x|h_i) := \frac{P(x|\pi)}{P(h_i|\pi)}.$$

Conversely, we may want to *begin* with players' beliefs over nodes $x \in h_i$ given h_i, and then ask whether or not a strategy is sequentially rational given those beliefs. For any player i, a *vector of beliefs* is a β_i in

$$\prod_{h_i \in H_i} \Delta(h_i);$$

given a profile π, β_i is called *weakly consistent with* π if given h_i and $x \in h_i$,

$$\beta_i(x|h_i) P(h_i|\pi) = P(x|\pi).$$

Weak consistency has no bite off the equilibrium path.[10] The profile π is *sequentially rational for i with beliefs $\beta_i(\cdot|h_i) \in \Delta(A(h_i))$ at h_i* if β_i is weakly consistent with π and

$$\pi_i(h_i) \in \arg\max_{d_i(h_i)} \left\{ \sum_{x \in h_i} \beta_i(x|h_i) u_i([\pi_i : d_i(h_i)], \pi_{-i}|x) \quad : \quad d_i(h_i) \in \Delta(A(h_i)) \right\}.$$

In games with perfect recall, equilibrium strategies are sequentially rational on the path of play.

[10] I.e., if $P(h_i|\pi) = 0$ then $P(x|\pi) = 0$ and $\beta_i(h_i)$ can be anything that adds up to one.

Proposition 2.3 *Let* π *be an equilibrium* (in the sense of Definition 8) *of a game with perfect recall. For any player* i *and history* h_i*, if* $P(h_i|\pi) > 0$ *then* π *is sequentially rational for player* i *at* h_i *with beliefs* $P_\pi(\cdot|h_i)$*.*

For instance, in Example 1.2, $\pi = ((b_1, y_1), w_2)$ is not sequentially rational for player 1 at the root node. What beliefs would make π sequentially rational for player 1 at the information set $h_1 = \{a_1w_2\ ,\ a_1x_2\}$? If player 1 believed that the conditional probability of player 2 having played w_2 given h_1 was p then

$$\begin{aligned} E_{\beta_1(h_1)}[u_1(y_1)|h_1] &= 3p + 2(1-p), \\ E_{\beta_1(h_1)}[u_1(z_1)|h_1] &= 4(1-p), \end{aligned}$$

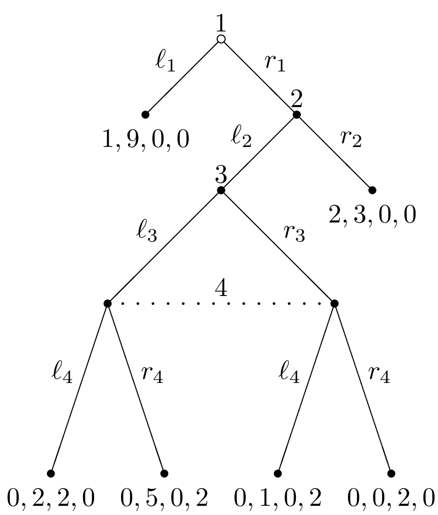
rendering y_1 conditionally optimal if $p \geq \frac{2}{5}$. It remains to check that such beliefs are weakly consistent. But since $P(a_1|\pi) = 0$ and *any* beliefs off the equilibrium path are weakly consistent, it follows that π is sequentially rational for player 1 at h_1 with beliefs $p \geq \frac{2}{5}$.

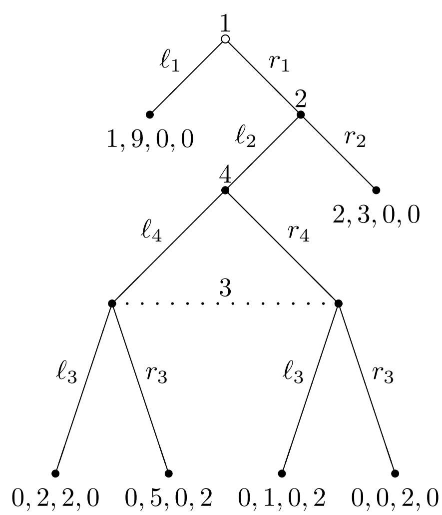
Definition 2.4 *A strategy profile* π *is* sequentially rational with beliefs β *if* π *is sequentially rational for every player* i *and every history* h_i *with beliefs* $\beta(\cdot|h_i)$*.*

2.3 Consistent Beliefs and Conditional Expectation

The inability of conditional expectation to prescribe beliefs off the equilibrium path is a problem. Consider the following related examples.

Example 2.5 *Consistency of sequentially rational profiles.*





The game on the left differs from that on the right in that the turns for players 3 and 4 are switched. In the left game, the profile $\pi = ([\ell_1], [\ell_2], \frac{1}{2}[\ell_3] + \frac{1}{2}[r_3], [\ell_4])$ is an equilibrium, yet there do not exist beliefs for player 3 that would make π sequentially rational for him given those beliefs.[11] Indeed, according to π, player 4 plays ℓ_4. Since sequential rationality only allows "deviations" by one player at a time, player 3 strictly prefers to play ℓ_3 rather than π_3 (since $2 > 1$). Thus, π_3 is not a best response.

This suggests that sequential rationality as defined in Section 2.2 *can* usefully refine the equilibrium notion of Definition 8, since π is an equilibrium that fails the test of sequential rationality. Alas, by making the apparently innocuous change from the left tree to the right tree, it turns out that not only is π an equilibrium of the right game, but it also *passes* the test of sequential rationality.

To see this, let x and y be the two nodes in player 3's information set, h_3, and let his beliefs be given by $\beta_3(x|h_3) = \beta_3(y|h_3) = \frac{1}{2}$. Since h_3 is reached with probability zero under π (player 1 plays ℓ_1), it is immediate that $\beta_3(\cdot|h_3)$ is weakly consistent with π. But given such beliefs, player 3 is indifferent between ℓ_3 and r_3 (each leads to an expected payoff of 1 miserable util). Therefore, it is a best response at h_3 to randomize between ℓ_3 and r_3. If player 3 plays $\frac{1}{2}[\ell_1] + \frac{1}{2}[r_1]$ then player 4 optimally plays ℓ_4, etc. The problem with this equilibrium is that it seems hard to justify player 3's beliefs when player 4 is playing the pure strategy ℓ_4.

Kreps and Wilson (1982) developed a stronger notion of consistency that eliminates both equilibria. Let Π° be the set of strategy profiles that assign positive probability to every node. For any such π° *every* history has positive probability, so is on the equilibrium path. Say that β° is *consistent* with π° if β° is weakly consistent with π°.[12]

The set of all consistent pairs (π°, β°) with $\pi^\circ \in \Pi^\circ$ is denoted by $\mathcal{C}^\circ$. To generally define beliefs β that are consistent with π even if some histories occur with zero probability under π, we will approximate π by a sequence of pairs $(\pi^\circ_n, \beta^\circ_n) \in \mathcal{C}^\circ$. Formally, for any profile π, a beliefs vector β is called *consistent* with π if there is a sequence $(\pi^\circ_n, \beta^\circ_n) \in \mathcal{C}^\circ$ such that

$$\begin{aligned} \pi_i(h_i)(a_i) &= \lim_{n\to\infty} \pi^\circ_{i,n}(h_i)(a_i), \qquad \forall i, h_i, a_i; \\ \beta_i(x|h_i) &= \lim_{n\to\infty} \beta^\circ_{i,n}(x|h_i), \qquad \forall i, h_i, x \in h_i. \end{aligned}$$

Since β° is consistent with π°, $\beta^\circ_{i,n}(x|h_i) = P(x|\pi^\circ_n)/P(h_i|\pi^\circ_n) \le 1$ for each n, so $\beta_i(x|h_i) \le 1$. Let $\mathcal{C}$ be the set of all such consistent pairs. (Clearly, $\mathcal{C} \supset \mathcal{C}^\circ$.) We are now ready to define a sequential equilibrium. Let π be any strategy profile and β any beliefs vector.

Definition 2.6 *The assessment (π, β) is called a* sequential equilibrium *if*

1. *π is sequentially rational with β,*
2. *β is consistent with π.*

Let's try some examples now.

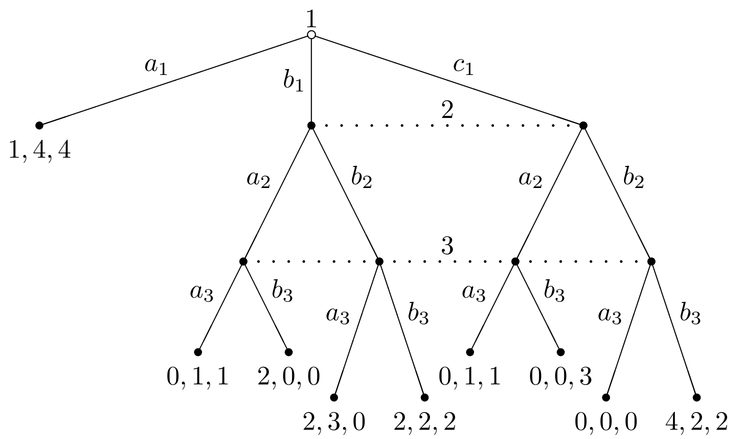
[11] By Proposition 10, it follows that $P(h_3|\pi) = 0$.

[12] Of course, since every node has positive probability under π°, β° is uniquely determined by Bayes' rule in the condition for weak consistency.

2.4 Examples of Sequential Equilibria

For either game in Example 2.5, the unique sequential equilibrium is $\pi = ([\ell_1], [\ell_2], \frac{1}{2}[\ell_3]+\frac{1}{2}[r_3], \frac{1}{2}[\ell_4]+\frac{1}{2}[r_4])$ together with beliefs $\beta_4 = (\frac{1}{2}, \frac{1}{2})$ in the left game and $\beta_3 = (\frac{1}{2}, \frac{1}{2})$ in the right game. The remaining players' beliefs are trivial since everyone else's information sets are singletons.

Example 2.7 *Another example of sequential equilibrium.*



Notice that $\pi = (a_1, a_2, a_3)$ is an equilibrium of this game. Let's find out whether or not there are beliefs β that turn (π, β) into a sequential equilibrium. Consider the following completely mixed perturbation π^ϵ of π.

$$\pi^\epsilon = ((1-\epsilon_0-\epsilon_1)[a_1] + \epsilon_0[b_1] + \epsilon_1[c_1], (1-\epsilon_2)[a_2] + \epsilon_2[b_2], (1-\epsilon_3)[a_3] + \epsilon_3[b_3]).$$

As $\epsilon \to 0$, i.e., as $\epsilon_i \to 0$ for every i, it is clear that $\pi^\epsilon \to \pi$. Applying Bayes' rule yields the following consistent beliefs: $\beta_1(\aleph|\aleph) = 1$ by a trivial application. Since mixing is independent and the only thing that could ever be learnt by Bayes' rule is that a_1 wasn't played,

$$\begin{aligned}
\beta_2^\epsilon(b_1|h_2) &= \frac{\epsilon_0}{\epsilon_0+\epsilon_1} = 1 - \beta_2^\epsilon(c_1|h_2), \\
\beta_3^\epsilon(b_1a_2|h_3) &= (1-\epsilon_2)\beta_2^\epsilon(b_1|h_2), \\
\beta_3^\epsilon(c_1a_2|h_3) &= (1-\epsilon_2)\beta_2^\epsilon(c_1|h_2), \\
\beta_3^\epsilon(b_1b_2|h_3) &= \epsilon_2\beta_2^\epsilon(b_1|h_2), \\
\beta_3^\epsilon(c_1b_2|h_3) &= \epsilon_2\beta_2^\epsilon(c_1|h_2).
\end{aligned}$$

As $\epsilon \to 0$, beliefs converge to

$$\begin{aligned}
\beta_2(b_1|h_2) &= p = 1 - \beta_2(c_1|h_2), \\
\beta_3(b_1a_2|h_3) &= p = 1 - \beta_3(c_1a_2|h_3), \\
\beta_3(b_1b_2|h_3) &= 0 = \beta_3(c_1b_2|h_3),
\end{aligned}$$

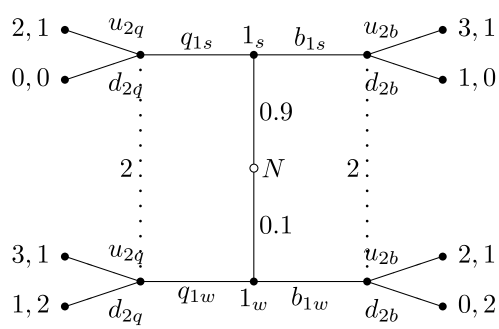
where $p = \lim \epsilon_0/(\epsilon_0 + \epsilon_1)$ could be any number between 0 and 1, depending on the rate of convergence of ϵ. For these beliefs to be consistent, it must be the case that

$$\begin{aligned} \beta_3(b_1a_2|h_3) + \beta_3(b_1b_2|h_3) &= \beta_2(b_1|h_2), \\ \beta_3(b_1a_2|h_3) + \beta_3(c_1a_2|h_3) &= \pi_2(a_2); \end{aligned}$$

both conditions are satisfied. Therefore β is consistent with π for any p. However, π is not sequentially rational with β. Indeed, if player 3 plays a_3 then with these beliefs player 2 would get an expected utility of $3(1-p)$ if he played b_2 and 1 if he played a_2. Hence $p \geq 2/3$ is necessary for sequential rationality. Likewise, player 3's expected utility from playing a_3 is 1 whereas the expected utility from playing b_3 is $3p$. Therefore, sequential rationality requires that $p \leq 1/3$, a contradiction.

It follows that (a_1, a_2, a_3) cannot be supported by beliefs that make it a sequential equilibrium. However, the profile (c_1, b_2, b_3) can be supported in sequential equilibrium (with the obvious beliefs).

Example 2.8 *The beer-quiche game.*



Let's show that the profile $\pi = ((b_{1s}, b_{1w}), (d_{2q}, u_{2b}))$ is a sequential equilibrium with some consistent beliefs. On the path of play, player 2 must have beliefs consistent with Bayes' rule. Therefore, $\beta_2(1s|b_1) = .9$, i.e., since both types of player 1 adopt the same strategy, player 2 doesn't learn from observing b_1. Therefore u_{2b} is sequentially rational for player 2, since player 2's expected payoff from playing u_{2b} is 1, which exceeds the payoff from playing d_{2b} ($2(0.1) = .2$ utils).

Suppose that player 2's beliefs about player 1's type conditional on observing quiche are given by $\beta_2(q_{1s}|q_1) = p$ and $\beta_2(q_{1w}|q_1) = 1 - p$. The action d_{2q} is sequentially rational given these beliefs if $2(1-p) \geq 1$. In words, given that player 2 observes that player 1 played q_1, playing d_2 will be a best response only when player 2 assigns sufficiently high probability to the event that player 1's type is 1_w. But if player 2 optimally plays d_{2q} then players 1_s and 1_w will have no incentive to deviate from b_1. Hence π is sequentially rational with these beliefs. Furthermore, such beliefs are consistent with π. Let $\pi_1^\epsilon = ((1-\epsilon_s)[b_{1s}] + \epsilon_s[q_{1s}], (1-\epsilon_w)[b_{1w}] + \epsilon_s[q_{1w}])$, with $\epsilon = (\epsilon_s, \epsilon_w) \to 0$. The beliefs $\beta_1^\epsilon(q_{1s}|q_1) = .9\epsilon_s/(.9\epsilon_s + .1\epsilon_w)$ are consistent with π^ϵ, and by picking the ratio ϵ_s/ϵ_w appropriately we can make them converge to any p in $[0, 1]$. We're done.

3 Self Confirming Equilibrium

First I will define the flora and fauna of self confirming equilibrium concepts and comment a little bit on them and then I'll introduce some examples.

3.1 Equilibrium Beliefs about Opponents' Play

Self-confirming equilibrium (SCE) is motivated by the idea that Nash equilibrium assumes players know their opponents' strategies. The object of SCE is to relax that assumption. To begin with, players are assumed to only know the structure of the extensive form, the distribution of nature's moves, and their own payoffs. The additional requirement for SCE is that players' beliefs about their opponents' play be correct on the equilibrium path, but not necessarily off the path. The space of players' beliefs is defined in terms of behavior strategy profiles. For any player i, i's beliefs about his opponents' play is any probability measure μ_i in

$$\Omega_i := \Delta\left(\Pi_{-i}\right).$$

Player i's beliefs μ_i reflect his uncertainty about opponents' play and therefore induce probabilistic beliefs on terminal nodes. If i plays the pure strategy c_i then his subjective probability that the terminal node z will be reached is simply the expected value of the conditional probability of reaching z given (c_i, π_{-i}), thus:

$$P_i(z|c_i, \mu_i) = E_{\mu_i}[P(z|c_i, \tilde{\pi}_{-i})] = \int_{\Pi_{-i}} P(z|c_i, \pi_{-i}) d\mu_i(\pi_{-i}).$$

Although behavior strategy profiles assume independent mixing between opponents and information sets, beliefs may be correlated. To illustrate, suppose that there are three players, and that players 2 and 3 simultaneously choose between the strategies U and D. Let

$$\mu_1(\pi_{-1}) = \begin{cases} \frac{1}{4} & \text{if } \pi_2(U) = \pi_3(U) = 1, \\ \frac{3}{4} & \text{if } \pi_2(U) = \pi_3(U) = \frac{1}{2}. \end{cases}$$

In this case, $P_1(U, U) = \frac{1}{4}1 + \frac{3}{4} \cdot \frac{1}{4} = 7/16$, $P_1(U, D) = P_1(D, U) = P_1(D, D) = 3/16$.

Player i's utility function is defined over terminal nodes. We define the induced expected utility over belief-action pairs (c_i, μ_i) as

$$u_i(c_i, \mu_i) = \sum_{z \in Z} u_i(z) P_i(z|c_i, \mu_i).$$

Given a mixed strategy $\sigma_j \in \Sigma_j$, let me denote by $\hat{\pi}_j(\cdot|\sigma_j)$ its behavioral representation. The event

$$E_i(h_j) := \{\pi_{-i} : \pi_j(h_j) = \hat{\pi}_j(h_j|\sigma_j)\}$$

is the set of behavior strategy profiles for i's opponents that are consistent with player j playing according to σ_j at information state h_j. Notice that i and j are fixed. Therefore, $E_i(h_j)$ contains all behavior profiles π_{-i} whose jth entry π_j coincides with $\hat{\pi}_j(h_j|\sigma_j)$ at history h_j. The strategy π_j could be anything outside h_j, as could the strategies of all other players.

We are now ready to define a Nash equilibrium and a SCE.

Definition 3.1 *A profile σ is a* Nash equilibrium *if for any $i \in I$ and $c_i \in C_i$ with $\sigma_i(c_i) > 0$ there exists $\mu_i \in \Omega_i$ such that*

1. $c_i \in \arg\max u_i(\cdot, \mu_i)$,
2. $\mu_i(E_i(h_j)) = 1$ *for every h_j and every $j \neq i$.*

The first condition requires that c_i be an optimum strategy given beliefs μ_i, and the second condition requires that every player's beliefs are correct regarding every behavior strategy of every opponent in the sense that $\mu_i(E_i(h_j)) = 1$, i.e., that the subjective probability for player i that player j is playing $\hat{\pi}_j(h_j|\sigma_j)$ at h_j is one for every h_j and every j.

Definition 3.2 *A profile σ is a* self-confirming equilibrium *if for any $i \in I$ and $c_i \in C_i$ with $\sigma_i(c_i) > 0$ there exists $\mu_i \in \Omega_i$ such that*

1. $c_i \in \arg\max u_i(\cdot, \mu_i)$,
2. $\mu_i(E_i(h_j)) = 1$ *for every h_j such that $P(h_j|\sigma) > 0$ and every $j \neq i$.*

The only difference between Nash equilibrium and SCE is that Nash equilibrium requires beliefs to be correct everywhere, whereas SCE only requires beliefs to be correct on information sets that have positive probability under σ, allowing for *wrong* beliefs about opponents' behavior strategies off the equilibrium path (i.e., on information sets h with $P(h|\sigma) = 0$).

Notice that in both definitions, potentially different μ_i's are allowed to justify different c_i's. In the definition of NE, this flexibility is vacuous, since beliefs—always having to be correct—must be unique (two different beliefs cannot both be correct). The flexibility matters once beliefs are allowed to be wrong. This suggests the following refinement, which requires that for every player i, if two strategies c_i and c_i' have positive probability of being played, then there must exist some beliefs that make both c_i and c_i' best responses *simultaneously*.

Definition 3.3 *A profile σ is a* unitary self-confirming equilibrium (USCE) *if given $i \in I$ there exists $\mu_i \in \Omega_i$ such that for any $c_i \in C_i$ with $\sigma_i(c_i) > 0$*

1. $c_i \in \arg\max u_i(\cdot, \mu_i)$,
2. $\mu_i(E_i(h_j)) = 1$ *for every h_j such that $P(h_j|\sigma) > 0$ and every $j \neq i$.*

Fudenberg and Levine (1993b) formalize the motivation of Nash equilibrium as the steady state of a learning process. They construct a dynamic environment where players repeatedly play an extensive-form game and, by playing, they learn about their opponents' strategies through Bayesian-updating. Intuitively, every player i considers trying different strategies to observe the play of his opponents, thus learning about his opponents' strategy profile, which is potentially beneficial to him. On the other hand i incurs an opportunity cost from experimenting by trying different strategies because some may lead to perhaps suboptimal strategies being played, depending on the actual profile of i's opponents' strategies.

Players have a common discount factor, δ. Formally, Fudenberg and Levine show that as $\delta \to 1$, the path of play that will arise in a steady state of the learning process is characterized by the set of Nash-equilibrium outcomes. Moreover, if $\delta < 1$, then players will generally play a self confirming equilibrium. The intuition behind this result is that as $\delta \to 1$, the opportunity cost of (optimum) experimentation tends to zero, implying that the benefits of learning exceed the zero cost, whence players will experiment until they know with probability 1 the strategy profile of their opponents. However, when $\delta < 1$, such opportunity cost is strictly positive, therefore it may be optimal to forego learning about opponents' strategies to some extent. Of course, on the equilibrium path, play by opponents is necessarily observed, so beliefs ought to be correct there.

So far, it may be argued, this intuitive story only captures the notion of *unitary* self confirming equilibrium. True enough. Fudenberg and Levine consider games with player *roles*. In such games, there is a large population of players who are randomly matched from $|I|$ many pools and whose roles $i \in I$ are allocated according to a player's pool once matched. Furthermore, there are many different kinds of player types in the population. (Some will be commitment types.) In this larger game, a player from pool i will play the game with opponents that are not necessarily the same players from period to period.

In the long run, a player from pool i will have played with all sorts of types of opponents. In particular, he will have observed the behavior of all sorts of commitment types. If the size of the population of commitment types is small relative to the "rational" types, a player may use the law of large numbers to distinguish between the play of commitment types and the play of rational types. Furthermore, a player in pool i will be able to discern the play of his "rational" opponents on all the information sets that could be reached by some profile of commitment and "rational" types of opponents when the player in pool i is playing some strategy c_i. Intuitively, a player "observes" the experimentation of his opponents (reflected in different types of opponents) but not of other types of players in the same pool. The concept of consistent self-confirming equilibrium naturally characterizes the equilibria that are consistent with this story. For any player $i \in I$ and pure strategy c_i, let

$$H(c_i) = \{h \in H : \exists c_{-i} \in C_{-i} \text{ s.t. } P(h|c) > 0\}.$$

The set $H(c_i)$ consists of those information sets where player i may be restricted to have correct beliefs about opponents' play if he gets to observe the experimentation of opponents. It is the histories that are reachable by some strategy profile of i's opponents, c_{-i}, when i plays c_i.

Definition 3.4 *A profile σ is a* consistent self-confirming equilibrium (CSCE) *if for any $i \in I$ and $c_i \in C_i$ with $\sigma_i(c_i) > 0$ there exists $\mu_i \in \Omega_i$ such that*

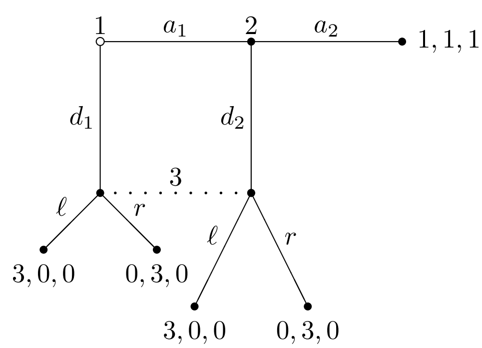
1. $c_i \in \arg\max u_i(\cdot, \mu_i)$,
2. $\mu_i(E_i(h_j)) = 1$ *for every* $h_j \in H(c_i)$ *and every* $j \neq i$.

It's about time to frame these ideas in specific games.

3.2 Examples of Self Confirming Equilibria

Here I present some examples to illustrate the different equilibrium concepts. The first example is perhaps the most important, originally presented by Fudenberg and Kreps (1988). This example shows that an SCE need not be Nash.

Example 3.5 *In the game below, (a_1, a_2) is a self-confirming equilibrium outcome.*



To show that (a_1, a_2) is the outcome of a SCE, suppose that player 1 believes that player 2 will play a_2 and player 3 will play r with probability 1. Then, player 1's best strategy is to play a_1. If player 2 believes that player 3 will play ℓ with probability 1, then player 2's best strategy will be to play a_2. Whatever player 3's strategy may actually be, he never gets to play it if players 1 and 2 play (a_1, a_2), whence the beliefs of players 1 and 2 about player 3's strategy never get disconfirmed. Since player 1's beliefs about player 2's strategy are correct, we arrive at a SCE.

However, this is not a Nash equilibrium, because players 1 and 2 have different beliefs about player 3's strategy and NE assumes that players have the same correct beliefs about everyone's strategy. Furthermore, there is no Nash equilibrium of this game that supports (a_1, a_2) as an outcome, for if players 1 and 2 both believed that player 3 would play ℓ with some probability p, then at least one of them would optimally play d_i.[13]

[13] If $p \leq \frac{2}{3}$ then player 2 would optimally play d_2, and if $p \geq \frac{1}{3}$ then player 1 would optimally play d_1.

Since players only have one information set each, this SCE is automatically unitary. Finally, notice that this is SCE is *not* consistent. To see this, notice that

$$H(a_1) = H(a_2) = \{h_1, h_2, h_3\},$$

where h_i is player i's only information set. Therefore, here the consistency refinement requires that both players 1 and 2 have correct beliefs about player 3's play. By the previous argument, (a_1, a_2) must therefore be suboptimal for at least one player.

One way to think about why this is happening is to notice that the game fails to have observed deviators. For any mixed-strategy profile σ, let

$$H(\sigma) = \{h \in H : P(h|\sigma) > 0\}$$

be the family of information sets that are reached with positive probability when players play σ.

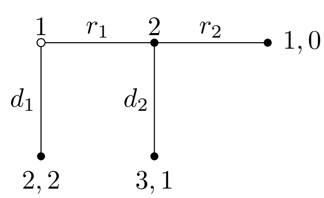
Definition 3.6 *A game has* observed deviators *if given a strategy profile c, any player i and possible deviation $d_i \neq c_i$, $h \in H(d_i, c_{-i}) \setminus H(c)$ implies that there is no d_{-i} such that $h \in H(c_i, d_{-i})$.*

This simply says that any history that may occur if i deviates from c_i to d_i cannot arise from player i playing c_i and some other player deviating. I.e., if player i deviates then somebody will find out that indeed it was player i who deviated. Games of perfect information and repeated games with observed actions have observed deviators. Also, two-player games with perfect recall satisfy this property, too. Fudenberg and Levine (1993a) prove the following classification of SCE in games with observed deviators.

Theorem 3.7 *In games with observed deviators, every self-confirming equilibrium is consistent.*

It follows that every SCE is a CSCE in two-player games with perfect recall.

Example 3.8 *The convexifying effect of self-confirming equilibria that are not unitary.*



In the game above, there are two Nash equilibrium *outcomes*: d_1 and (r_1, r_2). The mixed profile $(\frac{1}{2}[d_1] + \frac{1}{2}[r_1], r_2)$ is a SCE. However, it isn't unitary. Suppose that player 1 plays d_1 when he believes that player 2 will play d_2 with probability 1, and that player 1 plays r_1 when he believes that player 2 will play r_2 with probability 1. Then, when player 1 plays d_1, his beliefs about player 2's play are not disconfirmed. The mixed strategy by player 1 is somehow related more to random beliefs than to random behavior. It's as if a coin is flipped and if heads, player 1 will play d_1 *because* he believes that player 2 will play d_2, and if tails, player 1 will play r_1 *because* he believes that player 2 will play r_2.

Finally, it is easy to verify that the outcomes from non-unitary self confirming equilibria convexify the set of Nash equilibrium outcomes.

The last example reflects the possibility that correlated beliefs about opponents' play may lead to SCE that are not Nash outcomes.

Example 3.9 *Consider the following game with three players. Player 1 has the first move. He may choose one of four actions:* A, L_1, M_1, R_1. *If player 1 plays* A*, then the game ends and payoffs are* $(1, 0, 0)$. *If player 1 plays* L_1*,* M_1*, or* R_1*, then players 2 and 3 play the corresponding game:*

L_1:

	L_3	R_3
U_2	$4, 1, -1$	$0, -1, 1$
D_2	$0, -1, 1$	$-4, 1, -1$

M_1:

	L_3	R_3
U_2	$-4, 1, -1$	$0, -1, 1$
D_2	$0, -1, 1$	$4, 1, -1$

R_1:

	L_3	R_3
U_2	$0, 1, -1$	$3, -1, 1$
D_2	$3, -1, 1$	$0, 1, -1$

However, neither player 2 nor player 3 observes player 1's action.

In this example, the strategy A is a best response by player 1 to the correlated beliefs $P_1(U_2, L_3) = P_1(D_2, R_3) = \frac{1}{2}$, since for each possible strategy in $\{L_1, M_1, R_1\}$, its expected payoff to player 1 given these beliefs is zero, and the payoff from A is 1 util. However, A is never a best response to any strategy profile of players 2 and 3. For proof, suppose that player 2 plays U_2 with probability p_2 and that player 3 plays L_3 with probability p_3. In order for A to be a best response for player 1, it must be the case that

$$
\begin{aligned}
L_1 : &\quad 4p_2p_3 - 4(1-p_2)(1-p_3) \leq 1 \quad &\Rightarrow \quad & p_2 + p_3 \leq 5/4, \\
M_1 : &\quad -4p_2p_3 + 4(1-p_2)(1-p_3) \leq 1 \quad &\Rightarrow \quad & p_2 + p_3 \geq 3/4, \\
R_1 : &\quad 3p_2(1-p_3) + 3(1-p_2)p_3 \leq 1. & &
\end{aligned}
$$

If $p_2 \leq \frac{1}{2}$ then the lowest that p_3 could be whilst satisfying (R_1) is $p_3 = \frac{3}{4} - p_2$. Substituting this into (R_1) leads to the quadratic $2p_2^2 - \frac{3}{2}p_2 + \frac{3}{4}$, which is minimized with respect to p_2 when $p_2 = 3/8$. The implied value of p_3 is 3/8, too. But when $p_2 = p_3 = 3/8$, it is easy to check that the left-hand side of (R_1) is $30/64 > 1/3$, implying that (R_1) is violated. The proof if $p_2 > \frac{1}{2}$ is the same.

Let me underline that player 1 need not believe that players 2 and 3 correlate their play (indeed player 1 does not), it is precisely that player 1's *beliefs* are correlated.