A Foundation for Markov Equilibria in Infinite Horizon Perfect Information Games

V. Bhaskar, George J Mailath, and Stephen Morris

UCL, UPenn, and Princeton

EPSRC Symposium Workshop on Game Theory for finance, social and biological sciences (GAM)
April 15, 2010

link to paper

Introduction

- In the theory of repeated games, payoff irrelevant histories are used to create dynamic incentives.
- Many applications of game theory restrict attention to Markov equilibria — ruling this out.
- Possible justifications:
    - Markov equilibria are intuitive and simple;
    - Markov strategies focus on payoff-relevant dynamic incentive channels;
    - minors causes should have minor effects; and
    - coordination (bootstrapping) implicit in irrelevant history dependence sometimes is implausible.
    - Complexity costs imply simple strategies: in asynchronous games, complexity costs imply Markov equilibrium (Bhaskar/Vega-Redondo 02)
- But we still lack good foundations.

Main Result

Infinite Horizon Perfect Information Games

Strategy profile

- is purifiable if it is an equilibrium for some choice of independent continuous shocks at each decision node;
- has bounded recall if there is a uniform upper bound on the length of history that strategies depend on (for all but one player).

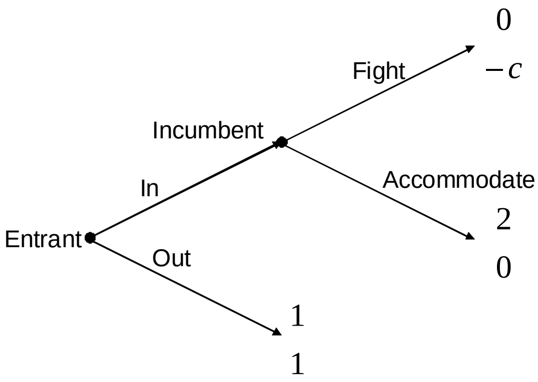
Theorem

All bounded recall purifiable equilibria are Markov.

Corollary

Repeated backward induction outcome is unique bounded recall purifiable equilibrium of infinitely repeated perfect information game.

Chain Store Game



- Long run incumbent; discount rate δ
- Short run entrants

Infinite History Equilibrium

- Trigger strategy equilibrium profile
    - Entrant: In only if the Incumbent has accommodated in the past. Otherwise, Out.
    - Incumbent: Fight if always fought in the past (and Entrant enters); otherwise Accommodate.
- Subgame Perfect Equilibrium for $\delta \geq c/(1+c)$.

Bounded Recall, Pure Strategy Equilibrium

- Entrant observes only last period outcome $h \in \{$Out,A,F$\}$.
- For generic δ, no pure strategy equilibrium with entry deterred (Ahn '97).
- Trigger strategy profile no longer an equilibrium because Incumbent has attractive one shot deviation to F if he chose A in the previous period (this pays under the above condition requiring sufficiently high δ):
  By choosing F, because next entrant only observes F, play reverts to Out.
- Same result if
  - entrant observes last K periods only.
  - if entrant observes infinite history but we restrict attention to strategies of entrant depending only on last period (or last K periods)

Bounded Recall, Mixed Strategy Equilibrium

Ahn's (mixed) strategy equilibrium profile:

- Entrant plays Out in first period and after first period plays
  - Out if $h =$Out or F; and
  - randomize with probability

$$\frac{c}{\delta(1+c)} \quad \text{on In and} \quad 1 - \frac{c}{\delta(1+c)} \quad \text{on Out}$$

  if $h =$A.
- Incumbent plays F with prob. $\frac{1}{2}$ independent of history.

Purification with 1-period Bounded Recall

- At each node, player moving gets additive payoff shock εz_i^t from first action, with z_i^t drawn with support $[0, 1]$ independently across players and nodes
- Entrant strategy: ρ_t : $\{$Out,A,F$\}$ $\times [0, 1] \rightarrow \Delta(\{In,Out\})$.
- Incumbent strategy: $\sigma_t : H^t \times [0, 1] \rightarrow \Delta(\{A,F\})$.
- NOTE: can show that players don't want to condition on earlier payoff shocks.

Purification with 1-period Bounded Recall

- Incumbent strategy σ_t (essentially) independent of h^{t-1} because
  - ρ_{t+1} is independent of h^{t-1}
  - noise implies (essentially) unique best response
- Entrant strategy ρ_t (essentially) independent of h^{t-1} because
  - σ_t is independent of h^{t-1}
  - noise implies (essentially) unique best response
- mixed strategy equilibrium not purifiable
- Always enter is only outcome path of purifiable bounded recall equilibria.

(Unperturbed) Infinite Game of Perfect Information

- Countable Set of Players, $\mathcal{N}$
- Countable Set of States, S
- Assignment of Player to States, $\iota : S \to \mathcal{N}$
  - i moves at states $S(i) = \{s \in S \mid \iota(s) = i\}$
- Countable Action Set, A
  - w.l.o.g. independent of S
- Transitions, $q : S \times A \to \Delta(S)$
  - $q(s' \mid s, a)$ is probability of state s' following state s when action a is played
- Initial distribution over states $q_0 \in \Delta(S)$
- Flow payoffs $u_i : S \times A \to \mathbb{R}$
- Discount rates $\delta_i \in [0, 1)$

Histories and Dynamic Payoffs

$\Gamma = \left\{S, \mathcal{N}, \iota, q, (\delta_i, u_i)_{i \in \mathcal{N}}\right\}$.

- period 0 history $H^0 = \{\varnothing\}$
- period $t \geq 1$ history $H^t = (S \times A)^t$
- all histories $H = \cup_{t=0}^{\infty} H^t$, with typical history $h \in H$
- payoffs given by

$$U_i\left((s_t, a_t)_{t=0}^{\infty}\right) = (1 - \delta_i) \sum_{t=0}^{\infty} \delta_i^t u_i(s_t, a_t).$$

Strategies

- strategy of player i,

$$b_i : H \times S(i) \to \Delta(A)$$

- strategy profile, $b = (b_i)_{i \in \mathcal{N}}$
- $V_i\,(b \mid h, s)$ is player i's expected continuation utility from the strategy profile b at the history (h, s):

$$V_i\,(b \mid h,s) = \sum_{a \in A} b_{\iota(s)}\,(a \mid h,s) \Big\{ (1-\delta_i)\,u_i\,(s,a) \\ + \delta_i \sum_{s' \in S} q\,(s' \mid s,a)\,V_i\,(b \mid (h,s,a)\,,s') \Big\}.$$

Strategies and Equilibrium

Definition

A strategy b_i is Markovian if for each $s \in S(i)$ and histories $h, h' \in H$ of the same length (i.e., $\tau(h) = \tau(h')$),

$$b_i(h, s) = b_i(h', s).$$

A Markovian strategy is stationary if the two histories can be of different lengths.

Definition

Strategy profile b is a subgame perfect Nash equilibrium (SPNE) if, for all $s \in S$, $h \in H$, and each $i \in \mathcal{N}$ and $b_i' \in B_i$,

$$V_i((b_i, b_{-i}) \mid h, s) \geq V_i((b_i', b_{-i}) \mid h, s).$$

Strategy profile b is a Markovian equilibrium if each b_i is Markovian and b is a SPNE.

Example I—asynchronous moves

	c_1	c_2	d
c_1	$11, 11$	$6, 9$	$-20, 20$
c_2	$9, 6$	$10, 10$	$-20, 20$
d	$20, -20$	$20, -20$	$0, 0$

- Player 1 moves in odd periods and player 2 in even periods (since time begins at $t = 0$, player 2 makes the first move).
- State and action sets are $S = A = \{c_1, c_2, d\}$ and the state encodes the action taken in the previous period (so $q(s' \mid s, a) = 1$ if $s' = a$ and 0 otherwise).
- Suppose the initial state is given by c_1.

There are two stationary pure strategy Markov equilibria:

1. Let $b^* : S \to A$ be the Markov strategy given by

$$b^*(s) = s.$$

   b^* is a perfect equilibrium for $\delta \in [\frac{1}{2}, \frac{20}{31}]$.

2. Let $b^\dagger : S \to A$ be the Markov strategy given by

$$b^\dagger(s) = \begin{cases} c_2, & \text{if } s = c_1, c_2, \\ d, & \text{if } s = d. \end{cases}$$

   $b^\dagger$ is a perfect equilibrium for $\delta \in [\frac{1}{2}, \frac{2}{3}]$.

1. There is a third mixed stationary Markov equilibrium: Let $b^\alpha : S \to \Delta(A)$ be the Markov strategy given by

$$b^\alpha(s) = \begin{cases} \alpha \circ c_1 + (1-\alpha) \circ c_2, & \text{if } s = c_1, \\ c_2, & \text{if } s = c_2, \\ d, & \text{if } s = d. \end{cases}$$

   This is an eq if $\alpha = (4\delta - 2)/[(5-\delta)\delta]$ for $\delta \in \left[\frac{1}{2}, \frac{2}{3}\right]$.
2. For any time t, the nonstationary Markov strategy specifying for periods before or at t, play according to b^*, and for periods after t, play according to b^α, is a Markov equilibrium for $\delta \in (\frac{1}{2}, \frac{2}{3})$. Not robust.
3. An outcome path of alternating c_1 and c_2 is the outcome path of a subgame perfect equilibrium that is not the outcome path of any Markov equilibrium. Not robust.

Example II—bargaining

1. In the initial period, a player $i \in \{1,2,3\}$ is selected randomly and uniformly to propose a coalition with one other player j, who can accept or reject. If j accepts, the game is over with payoffs:

coalition	1's payoff	2's payoff	3's payoff
$\{1,2\}$	9	3	0
$\{2,3\}$	0	9	3
$\{1,3\}$	3	0	9

2. If j rejects, play proceeds to the next period, with a new proposer randomly selected. If no coalition is formed, all players receive a payoff of 0.

1. For $\delta < 3/4$, unique Markov equilibrium and it is stationary and in pure strategies with immediate acceptance.
2. For $\delta > 3/4$, there is no Markov equilibrium in stationary pure strategies.
3. For $3/4 < \delta < \sqrt{3/4}$, there are two nonstationary pure strategy Markov equilibria. In one, offers are accepted in odd periods and rejected in even periods, while in the other offers are accepted in even periods and rejected in odd. Robust.
4. A non-Markov equilibrium when $3/4 < \delta < \sqrt{3/4}$: in the first period, if 1 is selected, then 1 chooses 3, who accepts. If 1 chooses 2, then 2 rejects, with play then following the Markov equilibrium with acceptance in odd periods (so there is acceptance next period). If 2 or 3 are selected, then play follows the Markov equilibrium with acceptance in even periods. Not robust.

The Perturbed Game

- Perturbations
  - Possible perturbations Z is full dimensional compact subset of $\mathbb{R}^{|A|}$.
  - $\Delta^* (Z)$ measures from strictly positive densities with support Z.
  - Fix $\varepsilon > 0$.
- At (h, s), for each i, $z_i \in Z$ is independently drawn under $\mu_i^s \in \Delta^*(Z)$.
  - Each player observes only own payoff shock.
  - Players' stage payoffs depend on current (s, a, z):

$$\tilde{u}_i (s, a, z) = u_i (s, a) + \varepsilon z_i^a .$$

- Denote the perturbed game by $\Gamma (\varepsilon, \mu)$.

Strategies and Equilibrium in Perturbed Game

- Each player knows history of his own past shocks, and so his behavior in principle can depend on this history.
- Since this history is private, to predict any player's behavior, other players form beliefs over this private history.
- But, conditional on the public history, player i's beliefs (about other players' private histories) are independent of player i's private payoff shocks.
- Consequently, any sequential best reply must be shock history independent: behavior at (h, s) can only depend on the payoff shock realized at (h, s), and is independent of earlier shocks.

Simple Characterization of Payoffs with Shock History Independent Strategies

- Shock history independent strategy (ignoring realization of z of measure 0) can be written as

$$\tilde{b}_i : H \times S(i) \times Z \to \Delta(A).$$

- If *all* players are following shock history independent strategies, we can recursively define value functions for a given strategy profile $\tilde{b}$ that do not depend on *any* payoff shock realizations:

$$V_i^*(\tilde{b} \mid h, s) = \int \sum_{a \in A} \tilde{b}_{\iota(s)}(a \mid h, s, z) \Bigg[(1 - \delta_i)\tilde{u}_i(s, a, z) + \delta_i \sum_{s' \in S} q(s' \mid s, a) V_i^*(\tilde{b} \mid (h, s, a), s') \Bigg] \mu^s(dz).$$

Don't need to worry about beliefs at unreached information sets

Lemma

Strategy profile $\widetilde{b}$ satisfies mutual sequential best responses (i.e., is a perfect Bayesian equilibrium) if and only if (i) each $\widetilde{b}_i$ is shock history independent; and (ii) for each h, $s \in S(i)$ and $\widetilde{b}'_i$,

$$V_i^*((\tilde{b}_i, \tilde{b}_{-i}) \mid h, s) \geq V_i^*((\tilde{b}'_i, \tilde{b}_{-i}) \mid h, s).$$

K-Recall Strategies

Definition

Shock history independent strategy $\widetilde{b}_i$ has K-recall if for all $s \in S(i)$, h, h' with $\tau(h) = \tau(h')$ and almost all z,

$$\widetilde{b}_i(h, s, z) = \widetilde{b}_i(h', s, z)$$

whenever

$$(s_k, a_k)_{k=t-K}^{t-1} = (s'_k, a'_k)_{k=t-K}^{t-1}.$$

If $\widetilde{b}_i$ has 0-recall, it is Markovian. If $\widetilde{b}_i$ has does not have K-recall for any finite K, it has infinite recall.

Key result

If one player does not have K-recall strategy, another player does not have (K+1)-recall strategy

Lemma

If $\widetilde{b}_i$ is a sequential best response to $\widetilde{b}_{-i}$ and does not have K-recall, then, for some $j \neq i$, $\widetilde{b}_j$ does not have $(K+1)$-recall.

Proof.

If $\widetilde{b}_i$ does not have K-recall, then there exist h and h' with $\tau(h) = \tau(h') = t \geq K$ and $s \in S(i)$ with

$$(s_k, a_k)_{k=t-K}^{t-1} = (s'_k, a'_k)_{k=t-K}^{t-1}$$

and

$$\widetilde{b}_i(h, s, z) \neq \widetilde{b}_i(h', s, z)$$

for a set z with positive measure.

Now suppose $\widetilde{b}_j$ has $(K+1)$-recall for all $j \neq i$. Player i's continuation value from choosing a at (h, s) or (h', s) is identical. Essentially unique best response implies essentially equal best responses. Contradiction.

A Corollary

Lemma

If $\widetilde{b}_i$ is a sequential best response to $\widetilde{b}_{-i}$ and does not have K-recall, then, for some $j \neq i$, $\widetilde{b}_j$ does not have $(K+1)$-recall.

Corollary

If $\widetilde{b}$ is a perfect Bayesian equilibrium of the perturbed game, then either $\widetilde{b}$ is Markovian or at least two players have infinite recall.

Back to Unperturbed Games: Defining Purifiability

- Fix strategy profile b of the unperturbed game.
- Shock history independent strategy sequence $\tilde{b}_i^k$ in the perturbed game converges to a strategy b_i if for each $h \in H$, $s \in S_i$ and $a \in A$,

$$\int \tilde{b}_i^k(a \mid h, s, z)\mu^s(dz) \rightarrow b_i(a \mid h, s)$$

Definition

Strategy profile b is purifiable if there exists $\mu : S \rightarrow \Delta^*(Z)$ and $\varepsilon^k \rightarrow 0$ such that there exists $\tilde{b}^k \rightarrow b$, with each $\tilde{b}^k$ a PBE of $\Gamma\left(\mu, \varepsilon^k\right)$.

Main Result

Theorem

If b is a purifiable SPNE in which no more than one player has infinite recall, then b is Markovian.

Note that stationarity is not an implication of purifiability and bounded recall.

Corollary

Suppose Γ is a finite perfect information game.
If Γ has a unique backward induction equilibrium, then the only purifiable equilibrium in which no more than one player has infinite recall is the infinite repetition of the backward induction equilibrium of Γ.
If Γ has multiple backward induction equilibria, if an equilibrium is purifiable and no more than one player has infinite recall, then it is an infinite sequence of history-independent specifications of the same backward induction equilibrium.

1. Purifiability

All Markovian equilibria are purifiable in our sense if we restrict attention to finite players and states.

Even with the flexibility of allowing the shock distribution μ to depend on both k and the target behavior profile, b, the argument is not trivial, since

1. future payoffs (including the contributions from the payoff shocks) affect current values and so the returns from different state transitions, and
2. perturbing actions results in both perturbed flow payoffs and perturbed transitions over states.

2. Existence of Markov Equilibria

- Duffie-Geanakoplos-MasColell-McLennan 94
- Escobar 08
- Doraszelski Escobar 08

3. Uniqueness of Markov Equilibria

- Yes in Repeated Perfect Information Games with Generic Payoffs
- Not in general.
- But this multiplicity of Markov equilibria does not allow us to sustain any additional outcomes.

4. Different notions of purification

Purification is used to mean different things:

1. when can we guarantee that there exists an essentially pure equilibrium in a game by adding noise to payoffs? Radner-Rosnethal 82.
    - this is trivial in our setting
2. we fix behavior in an unperturbed game and ask if there exists a sequence of equilibria of a sequence perturbed games that converge to the desired behavior...

4. Different notions of purification

Harsanyi showed (for static games) that (under some regularity conditions) every equilibrium was the limit of a sequence of equilibria in every sequence of perturbed games.

- Strategy profile b is Harsanyi purifiable if, for every $\mu : S \to M$ and $\varepsilon^k \to 0$, the $\left(\mu, \varepsilon^k\right)_{k=1}^{\infty}$ perturbed games have a sequence of strategy profiles $\tilde{b}^k$ converging to b, with $\tilde{b}_i^k$ a sequential best response to $\tilde{b}_{-i}^k$ in $\Gamma\left(\mu, \varepsilon^k\right)$ for each i.
- We conjecture that with additional regularity assumptions, Markovian equilibria will be Harsanyi purifiable
- Doraszelski Escobar 08 provide results of this type

5. Simultaneous Move Repeated Games

- Our argument does not work with simultaneous move games.
  - Our results do extend as follows: Fix a game with perfect information. In general, imperfections in the monitoring result in additional equilibria and make the game look like a simultaneous move game. But, if the monitoring imperfections are smaller than the purification shocks, then we again are led to Markov eq (take limits in the right order).
- Mailath-Samuelson 06 and Mailath-Olszewski 08 identify strict finite recall strategy profiles sufficient to prove folk theorems. Strictness ensures purifiability.

5. Simultaneous Move Repeated Games

- Belief-free strategies in recent work on repeated games (Piccione 02, Ely-Valimaki 02 and Ely-Horner-Olszewski 05)- like the Ahn mixed strategy equilibrium - rely on different mixed strategies at ”payoff equivalent” histories.
- Bhaskar-Mailath-Morris 08 show that the one period recall strategies of Ely Valimaki 02 are not purifiable via one period recall strategies in the perturbed game; however, they are purifiable via infinite recall strategies. The purifiability of such belief free strategies via finite recall strategies remains an open question (seems unlikely).

6. Endogenous Identification of Markovian Structure

- We constructed game so that states S capture everything payoff relevant.
- If there exists coarser state-labelling that is sufficient for payoff-relevance (a la Maskin-Tirole 01) our results would apply w.r.t. the coarser labelling (assuming shocks were measurable w.r.t. coarser labelling.