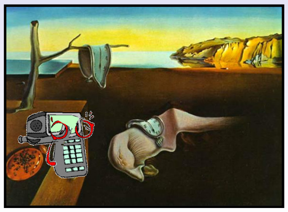
The Notion of Time and Imperfections in Games and Networks

Extensive Form Games, Repeated Games, Convergence Concepts in Normal Form Games, Trembling Hand Games, Noisy Observations



Model Timing Review

- When decisions are made also matters and different radios will likely make decisions at different time
- T_j – when radio j makes its adaptations
  - Generally assumed to be an infinite set
  - Assumed to occur at discrete time
    - Consistent with DSP implementation
- $T=T_1\cup T_2\cup\cdots\cup T_n$
- $t\in T$

Decision timing classes

- Synchronous
  - All at once
- Round-robin
  - One at a time in order
  - Used in a lot of analysis
- Random
  - One at a time in no order
- Asynchronous
  - Random subset at a time
  - Least overhead for a network

Motivating Extensive Form Games (1/2)

Models attempt to capture as much information as possible in as simple a format as possible. Often this requires abstracting away certain system details that are deemed unimportant to the analysis.

To this point, we have only considered systems where all decision-making occurs simultaneously and where action sets are defined independently of the decision making process.

Although many systems can be abstracted to fit these conventions, it is sometimes clumsy and if the analyst is not careful, may eliminate information that would alter the expected outcome.

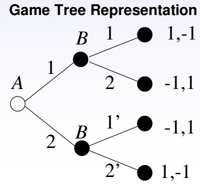
Extensive Form Game Components

Components

1. A set of players.
2. The actions available to each player at each decision moment (state).
3. A way of deciding who is the current decision maker.
4. Outcomes on the sequence of actions.
5. Preferences over all outcomes.

A Silly Jammer Avoidance Game

Game Tree Representation



Strategies for *A*

{1,2}

Strategies for *B*

{(1,1'),(1,2'), (2,1'),(2,2')}

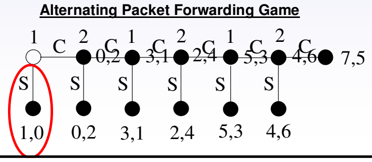
Strategic Form Equivalence

	1,1'	1,2'	2,1'	2,2'
1	1,-1	1,-1	-1,1	-1,1
2	-1,1	1,-1	-1,1	1,-1

Backwards Induction

- Concept
  - Reason backwards based on what each player would rationally play
  - Predicated on Sequential Rationality
  - Sequential Rationality – if starting at any decision point for a player in the game, his strategy from that point on represents a best response to the strategies of the other players
  - Subgame Perfect Nash Equilibrium is a key concept (not formally discussed today).

Alternating Packet Forwarding Game



Comments on Extensive Form Games

- Actions will generally not be directly observable
- However, likely that cognitive radios will build up histories
- Ability to apply backwards induction is predicated on knowing other radio's objectives, actions, observations and what they know they know…
  - Likely not practical
- Really the best choice for modeling notion of time when actions available to radios change with history

Repeated Games

- Same game is *repeated*
  - Indefinitely
  - Finitely
- Players consider discounted payoffs across multiple stages
  - Stage k
    $$\tilde{u}_i\left(a^k\right) = \delta^k u_i\left(a^k\right)$$
  - Expected value over all future stages
    $$\hat{u}_i\left(\left(a^k\right)\right) = \sum_{k=0}^{\infty} \delta^k u_i\left(a^k\right)$$

Stage 1

Γ	N	W
n	(9.6,9.6)	(3.2,21)
w	(21,3.2)	(7,7)

Stage 2

Γ	N	W
n	(9.6,9.6)	(3.2,21)
w	(21,3.2)	(7,7)

Stage k

Γ	N	W
n	(9.6,9.6)	(3.2,21)
w	(21,3.2)	(7,7)

Lesser Rationality: Myopic Processes

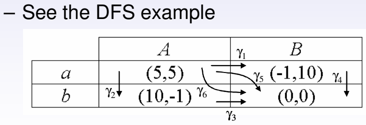
- Players have no knowledge about utility functions, or expectations about future play, typically can observe or infer current actions
- Best response dynamic – maximize individual performance presuming other players' actions are fixed
- Better response dynamic – improve individual performance presuming other players' actions are fixed
- Interesting convergence results can be established

Paths and Convergence

- *Path* [Monderer_96]
  - A path in Γ is a sequence $\gamma = (a^0, a^1, \ldots)$ such that for every $k \geq 1$ there exists a unique player such that the strategy combinations (a^{k-1}, a^k) differs in exactly one coordinate.
  - Equivalently, a path is a sequence of unilateral deviations. When discussing paths, we make use of the following conventions.
  - Each element of γ is called a *step.*
  - a^0 is referred to as the *initial* or *starting point* of γ.
  - Assuming γ is finite with m steps, a^m is called the *terminal point* or *ending point* of γ and say that γ has *length m.*
- *Cycle* [Voorneveld_96]
  - A finite path $\gamma = (a^0, a^1, \ldots, a^k)$ where $a^k = a^0$

Improvement Paths

- Improvement Path
  - A path $\gamma = (a^0, a^1, \ldots)$ where for all $k \geq 1$, $u_i(a^k) > u_i(a^{k-1})$ where i is the unique deviator at k
- Improvement Cycle
  - An improvement path that is also a cycle
  - See the DFS example



$\gamma_1 = ((a, A), (a, B))$	$\gamma_3 = ((b, A), (b, B))$	$\gamma_5 = (\gamma_1, (b, B))$
$\gamma_2 = ((a, A), (b, A))$	$\gamma_4 = ((a, B), (b, B))$	$\gamma_6 = (\gamma_1, (b, B))$

Convergence Properties

- Finite Improvement Property (FIP)
  - All improvement paths in a game are finite
- Weak Finite Improvement Property (weak FIP)
  - From every action tuple, there exists an improvement path that terminates in an NE.
- FIP implies weak FIP
- FIP implies lack of improvement cycles
- Weak FIP implies existence of an NE

Examples

Game with FIP

	A	B
a	1,-1	0,2
b	-1,1	2,2

Weak FIP but not FIP

	A	B	C
a	1,-1	-1,1	0,2
b	-1,1	1,-1	1,2
c	2,0	2,1	2,2

Implications of FIP and weak FIP

- Assumes radios are incapable of reasoning ahead and must react to internal states and current observations
- Unless the game model of a CRN has weak FIP, then no autonomously rational decision rule can be guaranteed to converge from all initial states under random and round-robin timing (Theorem 4.10 in dissertation).
- If the game model of a CRN has FIP, then ALL autonomously rational decision rules are guaranteed to converge from all initial states under random and round-robin timing.
  - And asynchronous timings, but not immediate from definition
- More insights possible by considering more refined classes of decision rules and timings

Decision Rules

Definition 4.10: *Best Response Dynamic*
A decision rule $d_i : A \rightarrow A_i$ is a best response dynamic if each adaptation would maximize the radio's utility if all other radios continued to implement the same waveforms, i.e., $d_i(a) \in \{b_i \in A_i : u_i(b_i, a_{-i}) \geq u_i(a_i, a_{-i}) \forall a_i \in A_i\}$

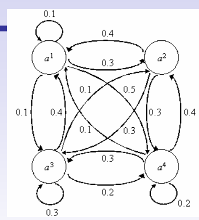
Definition 4.11: *Better Response Dynamic*
A decision rule $d_i : A \rightarrow A_i$ is a better response dynamic if each adaptation would improve the radio's utility if all other radios continued to implement the same waveforms, i.e., $d_i(a) \in \{b_i \in A_i : u_i(b_i, a_{-i}) > u_i(a_i, a_{-i})\}$.

Definition 4.13: *Friedman's Random Better Response* [Friedman_01]
Player i chooses an action from $A_i \backslash b_i$ where b_i is player i's current action according to a uniform random distribution. If the chosen action would improve the utility of player i, it is implemented, otherwise, the player continues to play b_i.

Definition 4.12: *Random Better Response Dynamic* (*)
A decision rule $d_i : A \rightarrow A_i$ is a random better response dynamic if for each $t_i \in T_i$, radio i chooses an action from A_i where each action has a nonzero probability of being chosen and implements the action if it would improve its utility.

Markov Chains

- Describes adaptations as probabilistic transitions between network states.
  - d is nondeterministic
- Sources of randomness:
  - Nondeterministic timing
  - Noise
- Frequently depicted as a weighted digraph or as a transition matrix

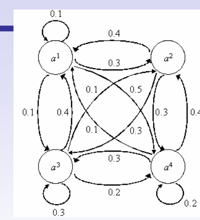


P =

	a^1	a^2	a^3	a^4
a^1	0.1	0.3	0.1	0.5
a^2	0.4	0.0	0.3	0.3
a^3	0.4	0.1	0.3	0.2
a^4	0.1	0.4	0.3	0.2

General Insights ([Stewart_94])

- Probability of occupying a state after two iterations.
  - Form **PP**.
  - Now entry p^{mn} in the m^{th} row and n^{th} column of **PP** represents the probability that system is in state a^n two iterations after being in state a^m.
- Consider P**k**.
  - Then entry p^{mn} in the m^{th} row and n^{th} column of represents the probability that system is in state a^n two iterations after being in state a^m.



P =

	a^1	a^2	a^3	a^4
a^1	0.1	0.3	0.1	0.5
a^2	0.4	0.0	0.3	0.3
a^3	0.4	0.1	0.3	0.2
a^4	0.1	0.4	0.3	0.2

Steady-states of Markov chains

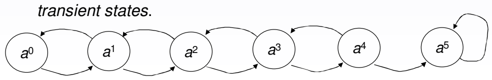
- May be inaccurate to consider a Markov chain to have a fixed point
  - Actually ok for absorbing Markov chains
- Stationary Distribution
  - A probability distribution such that $\boldsymbol{\pi}^*$ such that $\boldsymbol{\pi}^{*T}$ **P** $= \boldsymbol{\pi}^{*T}$ is said to be a stationary distribution for the Markov chain defined by **P**.
- Limiting distribution $\lim_{k \to \infty} \pi^{0T} \mathbf{P}^k$
  - Given initial distribution $\boldsymbol{\pi}^0$ and transition matrix **P**, the *limiting distribution* is the distribution that results from evaluating

Ergodic Markov Chain

- [Stewart_94] states that a Markov chain is ergodic if it is a Markov chain if it is a) *irreducible*, b) *positive recurrent*, and c) *aperiodic*.
- Easier to identify rule:
  - For some k **P**$^{\mathbf{k}}$ has only nonzero entries
- (Convergence, steady-state) If ergodic, then chain has a unique limiting stationary distribution.

Absorbing Markov Chains

- Absorbing state
  - Given a Markov chain with transition matrix **P**, a state a^m is said to be an absorbing state if $p^{mm}=1$.
- Absorbing Markov Chain
  - A Markov chain is said to be an *absorbing Markov chain* if
    - it has at least one absorbing state and
    - from every state in the Markov chain there exists a sequence of state transitions with nonzero probability that leads to an absorbing state. These nonabsorbing states are called *transient states*.



Absorbing Markov Chain Insights

- Canonical Form $\mathbf{P}' = \begin{bmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I}^{ab} \end{bmatrix}$
- Fundamental Matrix $\mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1}$
- Expected number of times that the system will pass through state a^m given that the system starts in state a^k.
  - n^{km}
- (Convergence Rate) Expected number of iterations before the system ends in an absorbing state starting in state a^m is given by t^m where 1 is a ones vector
  - **t=N1**
- (Final distribution) Probability of ending up in absorbing state a^m given that the system started in a^k is b^{km} where $\mathbf{B} = \mathbf{NR}$

Absorbing Markov Chains and Improvement Paths

- Sources of randomness
  - Timing (Random, Asynchronous)
  - Decision rule (random decision rule)
  - Corrupted observations
- An NE is an absorbing state for autonomously rational decision rules.
- Weak FIP implies that the game is an absorbing Markov chain as long as the NE terminating improvement path always has a nonzero probability of being implemented.
- This then allows us to characterize
  - convergence rate,
  - probability of ending up in a particular NE,
  - expected number of times a particular transient state will be visited

Connecting Markov models, improvement paths, and decision rules

- Suppose we need the path $\gamma = (a^0, a^1, \ldots a^m)$ for convergence by weak FIP.
- Must get right sequence of players and right sequence of adaptations.
- Friedman Random Better Response
  - Random or Asynchronous
    - Every sequence of players have a chance to occur
    - Random decision rule means that all improvements have a chance to be chosen
  - Synchronous not guaranteed

Γ	A	B
a	(1,1)	(0,0)
b	(0,0)	(1,1)

- Alternate random better response (chance of choosing same action)
  - Because of chance to choose same action, every sequence of players can result from every decision timing.
  - Because of random choice, every improvement path has a chance of occurring

Convergence Results (Finite Games)

	Timings			
Decision Rules	Round-Robin	Random	Synchronous	Asynchronous
Best Response	1,3	1,3	1	1,3
Exhaustive Better Response	3	3	-	3
Random Better Response[(a)]	1,2,3	1,2,3	1,2,3	1,2,3
Random Better Response[(b)]	1,3	1,2,3	1	1,2,3

(a) Definition 4.12, (b) Definition 4.13, 1. IESDS, 2. Weak FIP, 3. FIP

- If a decision rule converges under round-robin, random, or synchronous timing, then it also converges under asynchronous timing.
- Random better responses converge for the most decision timings and the most surveyed game conditions.
  - Implies that non-deterministic procedural cognitive radio implementations are a good approach if you don't know much about the network.

Impact of Noise

- Noise impacts the mapping from actions to outcomes, $f : A \to O$
- Same action tuple can lead to different outcomes
- Most noise encountered in wireless systems is theoretically unbounded.
- Implies that every outcome has a nonzero chance of being observed for a particular action tuple.
- Some outcomes are more likely to be observed than others (and some outcomes may have a very small chance of occurring)

DFS Example

- Consider a radio observing the spectral energy across the bands defined by the set C where each radio k is choosing its band of operation f_k.
- Noiseless observation of channel c_k

$$o_i(c_k) = \sum_{k \in N} g_{ki} p_k \theta(c_k, f_k)$$

- Noisy observation $\tilde{o}_i(c_k) = \sum_{k \in N} g_{ki} p_k \theta(c_k, f_k) + n_i(c_k, t)$
- If radio is attempting to minimize inband interference, then noise can lead a radio to believe that a band has lower or higher interference than it does

Trembling Hand ("Noise" in Games)

- Assumes players have a nonzero chance of making an error implementing their action.
  - Who has not accidentally handed over the wrong amount of cash at a restaurant?
  - Who has not accidentally written a "tpyo"?
- Related to errors in observation as erroneous observations cause errors in implementation (from an outside observer's perspective).

Noisy decision rules

- Noisy utility $\tilde{u}_i(a,t) = u_i(a) + n_i(a,t)$

Trembling Hand

Definition 4.20: *Friedman's Noisy Random Better Response* [Friedman_01]
Player i chooses an action $a_i \in A_i \backslash b_i$ where b_i is player i's current action according to a uniform random distribution. If $u_i(a_i,a_{-i}) > u_i(b_i,a_{-i})$, then a_i is implemented, however, if $u_i(a_i,a_{-i}) \leq u_i(b_i,a_{-i})$, then player i still switches to a_i with nonzero probability ρ.

Observation Errors

Definition 4.21: *Noisy Best Response Dynamic* (*)
A decision rule $\tilde{d}_i : A \times T \to A_i$ is a noisy best response dynamic if each adaptation would maximize the radio's noisy utility if all other radios continued to implement the same waveforms, i.e., $\tilde{d}_i(a) \in \{b_i \in A_i : \tilde{u}_i(b_i, a_{-i}, t) \geq \tilde{u}_i(a_i, a_{-i}, t) \forall a_i \in A_i\}$

Definition 4.22: *Noisy Better Response Dynamic* (*)
A decision rule $\tilde{d}_i : A \times T \to A_i$ is a noisy better response dynamic if each adaptation would improve the radio's utility if all other radios continued to implement the same waveforms, i.e., $\tilde{d}_i(a) \in \{b_i \in A_i : \tilde{u}_i(b_i, a_{-i}, t) > \tilde{u}_i(a_i, a_{-i}, t)\}$.

Definition 4.23: *Noisy Random Better Response Dynamic* (*)
A decision rule $\tilde{d}_i : A \times T \to A_i$ is a random better response dynamic if for each $t_i \in T_i$, radio i chooses an action from A_i with nonzero probability and implements the action if it would improve $\tilde{u}_i$.

Implications of noise

- For random timing, [Friedman] shows game with noisy random better response is an ergodic Markov chain.
- Likewise other observation based noisy decision rules are ergodic Markov chains
  - Unbounded noise implies chance of adapting (or not adapting) to any action
  - If coupled with random, synchronous, or asynchronous timings, then CRNs with corrupted observation can be modeled as ergodic Makov chains.
  - Not so for round-robin (violates aperiodicity)
- Somewhat disappointing
  - No real steady-state (though unique limiting stationary distribution)

DFS Example with three access points

- 3 access nodes, 3 channels, attempting to operate in band with least spectral energy.
- Constant power
- Link gain matrix

g_{ik}	1	2	3
1	1	0.5	0.1
2	0.5	1	0.3
3	0.1	0.3	1

- Noiseless observations

(f_1,f_1,f_1)	(f_1,f_1,f_2)	(f_1,f_2,f_1)	(f_1,f_2,f_2)	(f_2,f_1,f_1)	(f_2,f_1,f_2)	(f_2,f_2,f_1)	(f_2,f_2,f_2)	(f_2,f_1,f_1)
(0.6,0.8,0.4)	(0.5,0.5,0.0)	(0.1,0.0,0.1)	(0.0,0.3,0.3)	(0.0,0.3,0.3)	(0.1,0.0,0.1)	(0.5,0.5,0.0)	(0.6,0.8,0.4)	(0.0,0.3,0.3)

- Random timing

Trembling Hand

- Transition Matrix, p=0.1

P	(f_1,f_1,f_1)	(f_1,f_1,f_2)	(f_1,f_2,f_1)	(f_1,f_2,f_2)	(f_2,f_1,f_1)	(f_2,f_1,f_2)	(f_2,f_2,f_1)	(f_2,f_2,f_2)
(f_1,f_1,f_1)	0	1/3	1/3	0	1/3	0	0	0
(f_1,f_1,f_2)	1/30	3/10	0	1/3	0	1/3	0	0
(f_1,f_2,f_1)	1/30	0	9/10	1/30	0	0	1/30	0
(f_1,f_2,f_2)	0	1/30	1/3	3/5	0	0	0	1/30
(f_2,f_1,f_1)	1/30	0	0	0	3/5	1/3	1/30	0
(f_2,f_1,f_2)	0	1/30	0	0	1/30	9/10	0	1/30
(f_2,f_2,f_1)	0	0	1/3	0	1/3	0	3/10	1/30
(f_2,f_2,f_2)	0	0	0	1/3	0	1/3	1/3	0

- Limiting distribution

(f_1,f_1,f_1)	(f_1,f_1,f_2)	(f_1,f_2,f_1)	(f_1,f_2,f_2)	(f_2,f_1,f_1)	(f_2,f_1,f_2)	(f_2,f_2,f_1)	(f_2,f_2,f_2)
0.0161	0.0293	0.3846	0.0699	0.0699	0.3846	0.0293	0.0161

Noisy Best Response

- Transition Matrix, $\mathcal{N}(0,1)$ Gaussian Noise

P	(f_1,f_1,f_1)	(f_1,f_1,f_2)	(f_1,f_2,f_1)	(f_1,f_2,f_2)	(f_2,f_1,f_1)	(f_2,f_1,f_2)	(f_2,f_2,f_1)	(f_2,f_2,f_2)
(f_1,f_1,f_1)	0.3367	0.2038	0.2381	0	0.2214	0	0	0
(f_1,f_1,f_2)	0.1295	0.4813	0	0.1854	0	0.2038	0	0
(f_1,f_2,f_1)	0.0953	0	0.6273	0.1479	0	0	0.1295	0
(f_1,f_2,f_2)	0	0.1479	0.1854	0.5548	0	0	0	0.1119
(f_2,f_1,f_1)	0.1119	0	0	0	0.5548	0.1854	0.1479	0
(f_2,f_1,f_2)	0	0.1295	0	0	0.1479	0.6273	0	0.0953
(f_2,f_2,f_1)	0	0	0.2038	0	0.1854	0	0.4813	0.1295
(f_2,f_2,f_2)	0	0	0	0.2214	0	0.2381	0.2038	0.3367

- Limiting stationary distributions

	(f_1,f_1,f_1)	(f_1,f_1,f_2)	(f_1,f_2,f_1)	(f_1,f_2,f_2)	(f_2,f_1,f_1)	(f_2,f_1,f_2)	(f_2,f_2,f_1)	(f_2,f_2,f_2)
σ=1.00	0.0709	0.1120	0.1765	0.1406	0.1406	0.1765	0.1120	0.0709
σ=0.50	0.0540	0.1040	0.1984	0.1436	0.1436	0.1984	0.1040	0.0540
σ=0.10	0.0129	0.0647	0.2857	0.1366	0.1366	0.2857	0.0647	0.0129
σ=0.05	0.0033	0.0397	0.3387	0.1183	0.1183	0.3387	0.0397	0.0033
σ=0.01	0	0.002	0.46	0.038	0.038	0.46	0.002	0

Comment on Noise and Observations

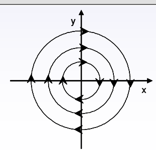
- Cardinality of goals makes a difference for cognitive radios
  - Probability of making an error is a function of the difference in utilities
  - With ordinal preferences, utility functions are just useful fictions
    - Might as well assume a trembling hand
- Unboundedness of noise implies that no state can be absorbing for most decision rules
- NE retains significant predictive power
  - While CRN is an ergodic Markov chain, NE (and the adjacent states) remain most likely states to visit
  - Stronger prediction with less noise
  - Also stronger when network has a Lyapunov function
  - Exception - elusive equilibria ([Hicks_04])

32/67

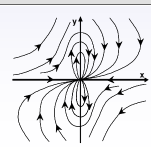
Stability

Definition 3.5: *Lyapunov stability*
We say that an action vector, a^*, is *Lyapunov stable* if for every $\varepsilon > 0$ there is a $\delta > 0$ such that for all $t \geq t^0$, $\|a(t^0), a^*\| < \delta \Rightarrow \|a(t), a^*\| < \varepsilon$ [8].

Definition 3.6: *Attractivity*
The action vector a^* is said to be *attractive* over the region $S \subset A$, $S = \{a \in A \mid \|a, a^*\| < M\}$, if given any $a(t_0) \in S$, the sequence $\{a(t)\}$ converges to a^* for $t \geq t_0$.



Stable, but not attractive



Attractive, but not stable

Cognitive Radio

33/67

Lyapunov's Direct Method

Theorem 3.3: *Lyapunov's Direct Method for Discrete Time Systems*
Given a recursion $a(t^{k+1}) = d^t(a(t^k))$ with fixed point a^*, we know that a^* is Lyapunov stable if there exists a continuous function (known as a **Lyapunov function***) that maps a neighborhood of a^* to the real numbers, i.e., $L: N(a^*) \to \mathbb{R}$, such that the following three conditions are satisfied:*

1) $L(a^*) = 0$
2) $L(a) > 0 \ \forall a \in N(a^*) \setminus a^*$
3) $\Delta L(a(t)) \equiv L[d^t(a(t))] - L(a(t)) \leq 0 \ \ \forall a \in N(a^*) \setminus a^*$

Further, if conditions 1-3 hold and

a) $N(a^*) = A$, *then a^* is globally Lyapunov stable;*
b) $\Delta L(a(t)) < 0 \ \forall a \in N(a^*) \setminus a^*$, *then a^* is asymptotically stable;*
c) $N(a^*) = A$ *and* $\Delta L(a(t)) < 0 \ \forall a \in N(a^*) \setminus a^*$, *then a^* is globally asymptotically stable.*

Left unanswered: where does *L* come from?
Can it be inferred from radio goals?

34/67

Summary

- Given a set of goals, an NE is a fixed point for all radios with those goals for all autonomously rational decision processes
- Traditional engineering analysis techniques can be applied in a game theoretic setting
  - Markov chains to improvement paths
- Network must have weak FIP for autonomously rational radios to converge
  - Weak FIP implies existence of absorbing Markov chain for many decision rules/timings
- In practical system, network has a theoretically nonzero chance of visiting every possible state (ergodicity), but does have unique limiting stationary distribution
  - Specific distribution function of decision rules, goals
  - Will be important to show Lyapunov stability
- Shortly, we'll cover potential games and supermodular games which can be shown to have FIP or weak FIP. Further potential games have a Lyapunov function!