01RMHNG–03RMHPF–01RMING
Network Dynamics
Week 10 — Part II
Learning Dynamics in Games

Giacomo Como and Fabio Fagnani
DISMA, Politecnico di Torino
{giacomo.como,fabio.fagnani}@polito.it
Torino, December 6, 2024

Outline

- Recap on (Potential) Games
- Finite Improvement Property
- Congestion Games
- Network Games

Learning dynamics in games

- Most economic theory relies on equilibrium analysis based on Nash equilibrium or its refinements.
- The traditional explanation for when and why equilibrium arises is that it results from the assumption of the rationality of the players, and that the structure of the game common shared knowledge.
- While there are situations where such elements are sufficient to determine the equilibrium (e.g., the prisoner dilemma by the dominant strategy technique), this is not the case in many situations where there are multiple Nash equilibria (e.g., coordination games)
- Below, we develop an alternative explanation why equilibrium arises as the long-run outcome of a learning process where players modify their action as time passes by.
- In contexts where the game is used as a modeling tool to solve multi-agent decision and optimization problems, such evolutionary dynamics can be interpreted as distributed algorithms.

Best Response dynamics (BR-dynamics)

$(\mathcal{V}, \{\mathcal{A}_i\}, \{u_i\})$ finite game, $\mathcal{X} = \prod_i \mathcal{A}_i$ configuration space

- $X(0) \in \mathcal{X}$ initial configuration
- At $t = 0, 1, \ldots$, player is chosen uniformly at random and activated;
- If player i activated at time t and $X(t) = x \in \mathcal{X}$, she computes

$$\mathcal{B}_i(x_{-i}) = \operatorname*{argmax}_{x_i \in \mathcal{A}} u_i(x_i, x_{-i})$$

  and then chooses her new action as

$$\mathbb{P}(X_i(t+1) = a \,|\, X(t) = x) = \begin{cases} \frac{1}{|\mathcal{B}_i(x_{-i})|} & \text{if } \; a \in \mathcal{B}_i(x_{-i}) \\ 0 & \text{otherwise} \end{cases}$$

  while $X_{-i}(t+1) = X_{-i}(t)$

$X(t)$ is a Markov chain on the configuration space $\mathcal{X}$.

Another graph in the configuration space

$(\mathcal{V}, \{\mathcal{A}_i\}, \{u_i\})$ game. $\mathcal{X} = \prod_i \mathcal{A}_i$ configuration space.

- Hypercube graph: $\mathcal{G}_h = (\mathcal{X}, \mathcal{E}_h)$ where

$$\mathcal{E}_h = \{(x, y) \,|\, \exists i \in \mathcal{V}, x_{-i} = y_{-i}, x_i \neq y_i\}$$

- Improvement graph (I-graph): $\mathcal{G}_I = (\mathcal{X}, \mathcal{E}_I)$

$$\mathcal{E}_I = \{(x, y) \,|\, \exists i \in \mathcal{V}, x_{-i} = y_{-i}, u_i(x) < u_i(y)\}$$

- Best Response graph (BR-graph): $\mathcal{G}_{BR} = (\mathcal{X}, \mathcal{E}_{BR})$

$$\mathcal{E}_{BR} = \{(x, y) \,|\, \exists i \in \mathcal{V}, x_{-i} = y_{-i}, y_i \in \mathcal{B}(x_{-i})\}$$

Analysis of the BR dynamics

$(\mathcal{V}, \{\mathcal{A}_i\}, \{u_i\})$ game. $\mathcal{X} = \prod_i \mathcal{A}_i$ configuration set.

$X(t)$ BR-dynamics on $\mathcal{X}$. $\mathcal{G}_{BR} = (\mathcal{X}, \mathcal{E}_{BR})$ BR-graph.

Analysis of the BR dynamics

$(\mathcal{V}, \{\mathcal{A}_i\}, \{u_i\})$ game. $\mathcal{X} = \prod_i \mathcal{A}_i$ configuration set.

$X(t)$ BR-dynamics on $\mathcal{X}$. $\mathcal{G}_{BR} = (\mathcal{X}, \mathcal{E}_{BR})$ BR-graph.

$$\text{For } x \neq y \quad P_{xy} = \begin{cases} \frac{1}{n} \frac{1}{|\mathcal{B}_i(x_{-i})|} & \text{if } (x, y) \in \mathcal{E}_{BR}, x_{-i} = y_{-i} \\ 0 & \text{if } (x, y) \notin \mathcal{E}_{BR} \end{cases}$$

Analysis of the BR dynamics

$(\mathcal{V}, \{\mathcal{A}_i\}, \{u_i\})$ game. $\mathcal{X} = \prod_i \mathcal{A}_i$ configuration set.

$X(t)$ BR-dynamics on $\mathcal{X}$. $\mathcal{G}_{BR} = (\mathcal{X}, \mathcal{E}_{BR})$ BR-graph.

$$\text{For } x \neq y \quad P_{xy} = \begin{cases} \frac{1}{n} \frac{1}{|\mathcal{B}_i(x_{-i})|} & \text{if } (x, y) \in \mathcal{E}_{BR}, x_{-i} = y_{-i} \\ 0 & \text{if } (x, y) \notin \mathcal{E}_{BR} \end{cases}$$

From the analysis of the BR-graph $\Rightarrow$ information on the asymptotics of the BR-dynamics

Asymptotics of the BR dynamics for potential games

- Theorem: In an ordinal potential game, every configuration x in a sink component of the BR-graph $\mathcal{G}_{BR}$ is a Nash equilibrium.

- Corollary: In an ordinal potential game, the BR dynamics $X(t)$ is such that, for every $X(0) \in \mathcal{X}$, there exists an almost surely finite random time $T \geq 0$ s.t. $X(t)$ is Nash for all $t \geq T$.

Proof of the Theorem

- Consider the BR-graph $\mathcal{G}_{BR} = (\mathcal{X}, \mathcal{E}_{BR})$. $\mathcal{N} \subseteq \mathcal{X}$ Nash equilibria.
- Notice that $(x, y) \in \mathcal{E}_{BR}$ implies $\Phi(y) \geq \Phi(x)$, namely the potential never decreases along the links of the transition graph.
- By contradiction: $\mathcal{W} \subseteq \mathcal{X}$ sink component and $\exists x \in \mathcal{W} \setminus \mathcal{N}$
- Then, $\exists y \in \mathcal{X}$ such that $(x, y) \in \mathcal{E}_{BR}$ and $\Phi(y) > \Phi(x)$.
- Since the potential never decreases, it means that there can not be a walk from y back to x. Hence, $y \notin \mathcal{W}$
- $\mathcal{W}$ not a sink component! $\Rightarrow$ Contradiction. ■

More questions on the BR dynamics for potential games

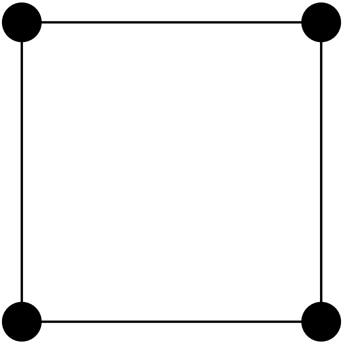
- What is the behavior of the BR dynamics within the set of Nash equilibria?
- Does it converge to the maximum points of the potential?
- Everything boils down to analyze the structure of the BR-graph $\mathcal{G}_{BR} = (\mathcal{X}, \mathcal{E}_{BR})$

More questions on the BR dynamics for potential games

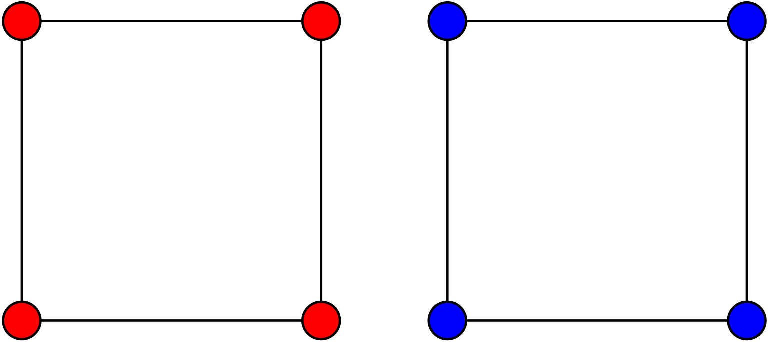
- What is the behavior of the BR dynamics within the set of Nash equilibria?
- Does it converge to the maximum points of the potential?
- Everything boils down to analyze the structure of the BR-graph $\mathcal{G}_{BR} = (\mathcal{X}, \mathcal{E}_{BR})$

- IMPORTANT: when we consider a graphical game $(\mathcal{V}, \{\mathcal{A}_i\}, \{u_i\})$ over a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, it is important not to confuse this graph with the BR-graph $\mathcal{G}_{BR}$ that is a graph on the configuration space $\mathcal{X}$!

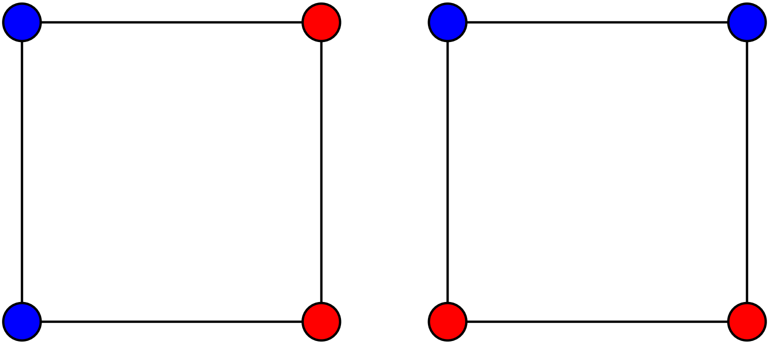
Majority game on the 4-cycle



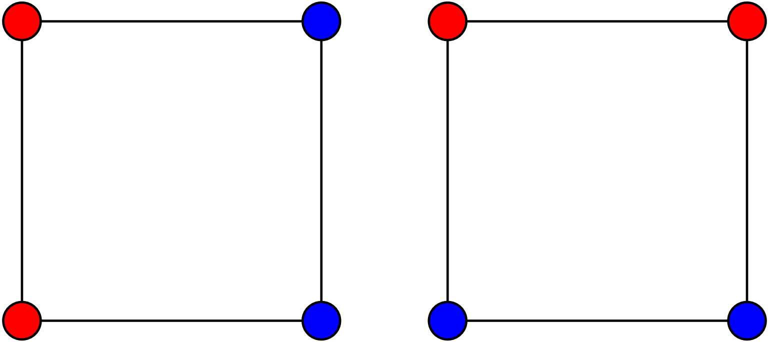
Majority game on the 4-cycle: Nash



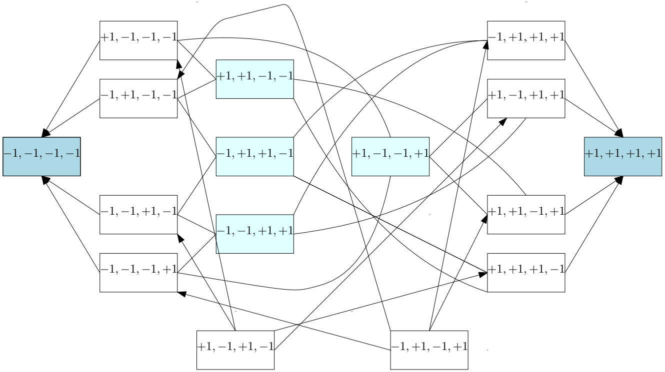
Majority game on the 4-cycle: Nash



Majority game on the 4-cycle: Nash



BR-graph of the Majority game over the 4-cycle



Nash (not strict)

Strict Nash, absorbing state

Behavior of the BR dynamics in potential games

Behavior of the BR dynamics in potential games

- There may be Nash equilibria that are not in sink components of the BR-graph.

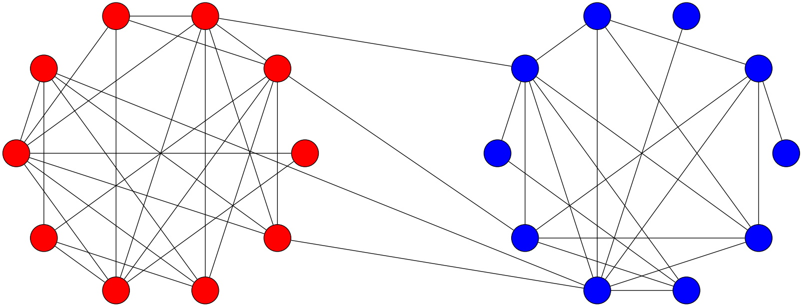
Behavior of the BR dynamics in potential games

- There may be Nash equilibria that are not in sink components of the BR-graph.
- The BR dynamics may enter and leave the set $\mathcal{N}$ of Nash equilibria.

Behavior of the BR dynamics in potential games

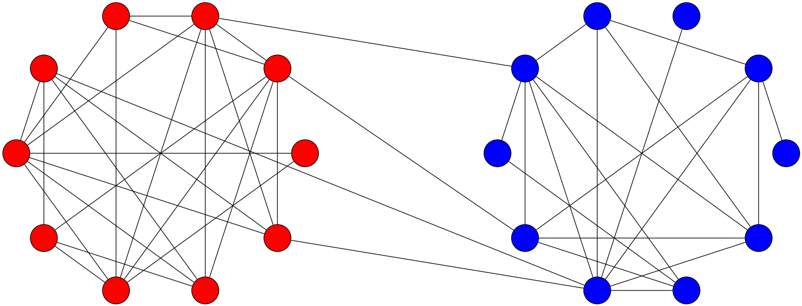
- There may be Nash equilibria that are not in sink components of the BR-graph.
- The BR dynamics may enter and leave the set $\mathcal{N}$ of Nash equilibria.
- The Nash equilibria that compose the sink components in general do not coincide with the maxima of the potential.

Behavior of the BR dynamics in potential games



This is a strict Nash equilibrium for the majority game, but it is not a maximum point of the potential.

Behavior of the BR dynamics in potential games

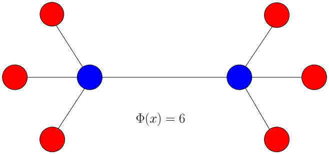


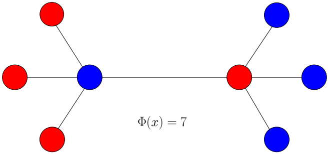
This is a strict Nash equilibrium for the majority game, but it is not a maximum point of the potential.

BR dynamics can get stuck in 'local maxima' and not converge to the global maxima of the potential!

Behavior of the BR dynamics in potential games

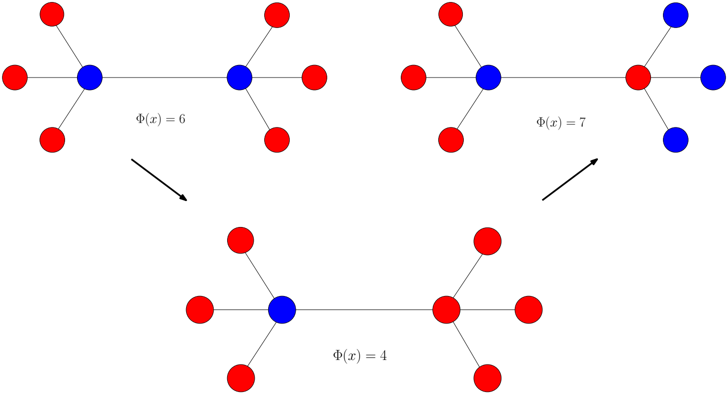
Two strict Nash equilibria for the network anticoordination game





Behavior of the BR dynamics in potential games

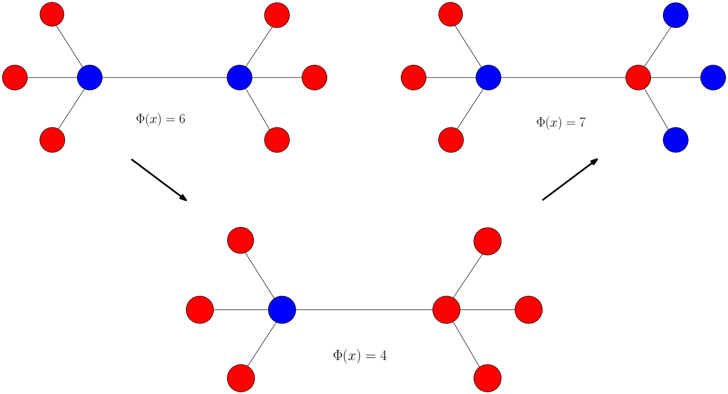
Two strict Nash equilibria for the network anticoordination game



The system needs to go through lower potential configurations in order to finally reach a maximum.

Behavior of the BR dynamics in potential games

Two strict Nash equilibria for the network anticoordination game



The system needs to go through lower potential configurations in order to finally reach a maximum. BR dynamics does not allow it!

Noisy Best Response dynamics (log-linear learning)

$(\mathcal{V}, \{\mathcal{A}_i\}, \{u_i\})$ game. $\mathcal{X} = \prod_i \mathcal{A}_i$ configuration set.

- $X(0) \in \mathcal{X}$ initial configuration
- At every discrete time instant, a player is chosen uniformly at random and becomes active;
- If node i becomes active at time t and the system is in configuration $X(t) = x \in \mathcal{X}$, he chooses its new action with probability

  $$\mathbb{P}(X_i(t+1) = a \,|\, X(t) = x) = \frac{1}{Z_i(x)} e^{\beta u_i(a, x_{-i})}$$

  where

  $$Z_i(x) = \sum_{a' \in \mathcal{A}_i} e^{\beta u_i(a', x_{-i})}$$

- $X_j(t+1) = X_j(t)$ for all $j \neq i$

Noisy Best Response dynamics (log-linear learning)

The NBR $X(t)$ is a Markov chain on $\mathcal{X}$.

The transition graph $G_{NBR} = (\mathcal{X}, \mathcal{E}_{NBR})$ is the *hypercube graph* (with self-loops):

$$x \neq y \in \mathcal{X},\, (x, y) \in \mathcal{E}_{NBR} \ \Leftrightarrow \exists i \in \mathcal{V} \, \text{s.t.}\, x_{-i} = y_{-i},\; x_i \neq y_i$$

The transition matrix is

$$\text{For}\, x \neq y \quad P_{xy} = \begin{cases} \frac{1}{n}\frac{1}{Z_i(x)} e^{\beta u_i(y_i, x_{-i})} & \text{if}\, (x, y) \in \mathcal{E}_{NBR}, x_{-i} = y_{-i} \\ 0 & \text{if}\, (x, y) \notin \mathcal{E}_{NBR} \end{cases}$$

Noisy Best Response dynamics (log-linear learning)

The NBR $X(t)$ is a Markov chain on $\mathcal{X}$.

The transition graph $G_{NBR} = (\mathcal{X}, \mathcal{E}_{NBR})$ is the *hypercube graph* (with self-loops):

$$x \neq y \in \mathcal{X},\, (x, y) \in \mathcal{E}_{NBR} \;\Leftrightarrow \exists i \in \mathcal{V}\,\text{s.t.}\; x_{-i} = y_{-i},\; x_i \neq y_i$$

The transition matrix is

$$\text{For}\, x \neq y \quad P_{xy} = \begin{cases} \frac{1}{n}\frac{1}{Z_i(x)} e^{\beta u_i(y_i, x_{-i})} & \text{if}\, (x, y) \in \mathcal{E}_{NBR}, x_{-i} = y_{-i} \\ 0 & \text{if}\, (x, y) \notin \mathcal{E}_{NBR} \end{cases}$$

- $\beta = 0$: NBR is a pure random walk on G_{NBR}
- $\beta \to +\infty$: NBR $\to$ BR.

Behavior of the NBR dynamics

Theorem

Consider an exact potential game $(\mathcal{V}, \{\mathcal{A}_i\}, \{u_i\})$ *with a potential* $\Phi : \mathcal{X} \to \mathbb{R}$*. Then, the NBR dynamics* $X(t)$ *is a reversible ergodic Markov chain whose unique equilibrium distribution is given by the Gibbs measure:*

$$\pi_x = \frac{e^{\beta\Phi(x)}}{\sum_{y \in \mathcal{X}} e^{\beta\Phi(y)}}$$

Moreover, considering that π *depends on* β*:*

$$\lim_{\beta \to +\infty} \pi = \textit{ uniform distribution over } \operatorname{argmax} \Phi$$

The proof

- The graph $\mathcal{G}^{\mathrm{NBR}}$ is strongly connected
- From the general theory, we know that the NBR $X(t)$ is ergodic: it possesses just one invariant distribution π
- To prove reversibility, we must find a positive vector ρ such that the balancing equation is satisfied:

$$\rho_x P_{xy} = \rho_y P_{yx} \quad \forall x, y$$

- Balancing equation trivially true if $x = y$ or if $P_{xy} = P_{yx} = 0$.

The proof

- Assume now that $x, y \in \mathcal{X}$ are such that $x_{-i} = y_{-i}$ and $x_i \neq y_i$:

$$\frac{P_{xy}}{P_{yx}} = \frac{\frac{1}{n}\frac{1}{Z_i(x)} e^{\beta u_i(y_i, x_{-i})}}{\frac{1}{n}\frac{1}{Z_i(y)} e^{\beta u_i(x_i, y_{-i})}} = e^{\beta[u_i(y_i, x_{-i}) - u_i(y_i, y_{-i})]} = e^{\beta[\Phi(x) - \Phi(y)]} = \frac{e^{\beta\Phi(x)}}{e^{\beta\Phi(y)}}$$

Hence,

$$e^{\beta\Phi(x)} P_{xy} = e^{\beta\Phi(y)} P_{yx}$$

- This says that $X(t)$ is reversible and that the unique equilibrium distribution is given by

$$\pi_x = \frac{e^{\beta\Phi(x)}}{\sum_{y\in\mathcal{X}} e^{\beta\Phi(y)}}$$

- $$\lim_{\beta\to+\infty} \frac{e^{\beta\Phi(x)}}{\sum_{y\in\mathcal{X}} e^{\beta\Phi(y)}} = \begin{cases} \dfrac{1}{|\operatorname{argmax}\Phi|} & \text{if } x \in \operatorname{argmax}\Phi \\ 0 & \text{otherwise} \end{cases}$$

Using the NBR dynamics

The NBR dynamics $X(t)$ can be effectively used to find the maxima of the potential:

$$\begin{aligned}
\lim_{\beta\to+\infty}\lim_{t\to+\infty}\mathbb{P}(X(t)=x) &= \lim_{\beta\to+\infty}\pi_x \\
&= \begin{cases} \dfrac{1}{|\operatorname{argmax}\Phi|} & \text{if } x \in \operatorname{argmax}\Phi \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

- If β is large, $X(t)$ will be, after a transient, with high probability on a maximum of the potential.
- As time passes by, $X(t)$ approximately samples uniformly on $\operatorname{argmax}\Phi$.

NBR for coloring graphs

$\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ undirected graph. Is $\mathcal{G}$ k-colorable?

- k-coloring game: $\mathcal{A} = \{1, \ldots, k\}$ colors, $\mathcal{X} = \mathcal{A}^{\mathcal{V}}$
- $u_i(x_i; x_{-i}) = -\sum_j W_{ij} \mathbb{1}_{\{x_i = x_j\}}$
- Potential: $\Phi(x) = -\frac{1}{2} \sum_{i,j} W_{ij} \mathbb{1}_{\{x_i = x_j\}}$
- $\max \Phi(x) = 0$ if and only if $\mathcal{G}$ is k-colorable
- In this case $\operatorname{argmax} \Phi$ are exactly the coloring with k colors
- NBR can be used as an approximate algorithm to find k-colorings of a graph