Best-response dynamics (with applications to distributed protocols and mechanisms)[*]

Lecturer: Paolo Penna

February 18, 2014

1 Warm up

Which of these three games have **pure Nash equilibria (PNE)**?

	H	T
H	1, -1	-1, 1
T	-1, 1	1, -1

Matching Pennies

	B	S
B	2, 1	0, 0
S	0, 0	1, 2

Battle of Sexes

[*]The material of this lecture is taken from Nisan et al. (2011a) where you can find several other applications of best-response mechanisms. There you have a more precise, extensive, and formal description of best-response mechanisms, plus further pointers into the literature.

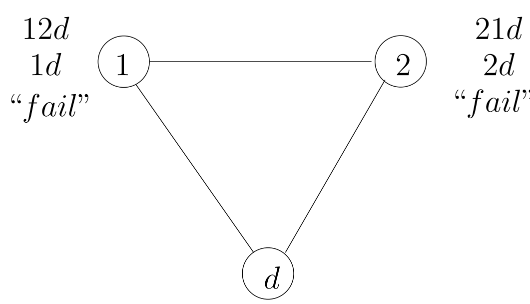
	A	S
A	-2, -2	0, -3
S	0, -3	-1, -1

Prisoners' Dilemma

> **(Synchronous) Best-Response Dynamics**: Players play their best response infinitely many times, one by one in a fixed order (round robin).

What happens for the three games above?

Example 1 Two nodes, 1 and 2, want to send traffic to another destination node d. Their strategy is to choose the *next hop* the traffic is sent to (one of the neighbors). The following picture shows the physical network and the preferences of each node (which path to use) near the corresponding node:



Each node prefers to reach d via the other node, but if they both send their own traffic to each other they fail (which is the least preferable option for both). ■

> **Question:** What happens if the two nodes move (play) always simultaneously? What happens if node 1 plays "1 → 2" at each step (while the other node plays best-response)?

Best Response:

1. No convergence in *asynchronous* settings.
2. Not incentive compatible.

For which games this does not happen?

Asynchronous Best-Response Dynamics: At each step an adversary activates an arbitrary subset of players who best respond to the current profile (the adversary also chooses a starting strategy profile). The adversary must activate each player an infinite number of times.

The choice of the adversary and the "response strategies" of each player determine an infinite sequence

$$s^0 \Longrightarrow s^1 \Longrightarrow \cdots s^t \Longrightarrow \cdots$$

If the game converges (after finitely many steps T we have $s^T = s^{T+1} = s^{T+2} = \cdots$) then the utility of each player i is $u_i(s^T)$. If the game keeps "oscillating" then we consider an upper bound on what the player can get (the worst case for us and the best for the player) that is $\limsup_{t\to\infty} u_i(s^t)$.

$$\begin{array}{rcl} \text{Base game } G & \Longrightarrow & \text{Repeated game } G^* \\ s_i \in S_i & & \text{response strategy } R_i() \in S_i \\ u_i(s) & & \text{total utility } \Gamma_i := \limsup\limits_{t\to\infty} u_i(s^t) \end{array}$$

Definition 2 *Best-response are* ***incentive compatible*** *for G if repeated best-responding is a Nash equilibrium for the repeated game G^*, that is, for every i*

$$\Gamma_i \geq \Gamma'_i$$

where Γ_i i the total utility when all players best respond and Γ'_i is the total utility when all but i best respond (starting from the same initial profile s^0 and applying the same activation sequence).

2 "Nice" Games

Consider this game (with a unique PNE):

		Player 2	
		A	B
Player 1	A	2, 1	0, 0
	B	3, 0	1, 2

Best response works as follows

$$(A,A) \xRightarrow{\textit{Player } 1} (B,A) \xRightarrow{\textit{Player } 2} (B,B) \xRightarrow{\textit{Player } 1} (B,B) \xRightarrow{\textit{Player } 2} (B,B) \cdots \Longrightarrow (B,B)$$

Player 1 improves if he/she does not best response (keep playing A):

$$(A,A) \xRightarrow{\textit{Player } 1} (A,A) \xRightarrow{\textit{Player } 2} (A,A) \xRightarrow{\textit{Player } 1} (A,A) \Longrightarrow \cdots \Longrightarrow (A,A)$$

Convergence but no incentive compatibility

Exercise 1 For the following game

		Player 2	
		A	B
Player 1	A	1, 1	1, 1
	B	1, 1	1, 1

find best response strategies that *never converge* (keep oscillating between different profiles). Find other best response strategies for which we *do have convergence*. ■

Two intuitions/ideas:

1. Introduce tie breaking rule.
2. Eliminate "useless" strategies.

2.1 Convergence

Consider this game

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

Exercise 2 *Prove that for this game best-response dynamics converge to a unique PNE.*

Note that in the previous game no strategy is dominant and no strategy is dominated. Strategy C satisfies the following (weaker) definition:

Definition 3 (never best response (NBR)) *A strategy $s_i \in S_i$ is a* never best response (for tie breaking rule $\prec$) *if there is always another strategy that gives a better payoff or that gives the same payoff but is better w.r.t. to this tie breaking rule: for all s_{-i} there exists $s'_i \in S_i$ such that one of these holds*

1. $u_i(s_i, s_{-i}) < u_i(s'_i, s_{-i})$ *or*
2. $u_i(s_i, s_{-i}) = u_i(s'_i, s_{-i})$ *and* $s_i \prec_i s'_i$.

The following condition is enough to guarantee convergence:

Definition 4 (NBR-solvable) *A game G is* NBR-solvable *if iteratively eliminating NBR strategies results in a game with one strategy per player. That is, there exists a tie breaking rule $\prec$, sequence $p_1, \ldots, p_\ell$ of players, and a corresponding sequence of subsets of strategies $E_1, \ldots, E_\ell$ such that:*

1. *Initially $G_0 = G$ and $G_i + 1$ is the game obtained from G_i by removing the strategies E_i of player p_i;*
2. *Strategies E_i are NBR for $\prec$ in the game G_{i-1}.*

3. *The final game G_ℓ has one strategy for each player (this unique profile is thus a PNE for G).*

A sequence of players and of strategies as above is called an elimination sequence *for the game G.*

Exercise 3 *Prove that the game described at the beginning of this section is NBR-solvable. Provide also a bound on the parameter ℓ.*

Lemma 5 (rounds vs subgames) *Let $p_1, \ldots, p_\ell$ be the players of any elimination sequence for the game under consideration. Suppose that players $p_1, \ldots, p_k$ always best respond (according to the prescribed tie breaking rule $\prec$). Then, for any initial profile and for any activation sequence, every profile after the k^{th} round is a profile in the subgame G_k.*

Before proving the lemma we observe that it implies convergence:

Theorem 6 (convergence) *For NBR-solvable games best response (according to the prescribed tie breaking rule $\prec$) converge even in the asynchronous case.*

Proof. Take $k = \ell$ and observe that G_ℓ contains only one profile. $\square$

Proof of Lemma 5. Denote by $round_j$ the last time step of the j^{th} round in the activation sequence. Obviously for any t we have $s^t \in G_0 = G$. Now consider $t \geq round_1$ and observe that, since player p_1 has been activated at least once the corresponding strategy satisfies [1]

$$s^t_{p_1} \notin E_1$$

which is equivalent to $s^t \in G_1$ for all $t \geq round_1$.

To prove the analogous for player p_2 we observe that, in the 2^{nd} round player p_2 is activated and, since $s^t \in G^1$ and since p_2 plays best response, for $t \geq round_2$ we have $s^t_{p_2} \notin E_2$. Since we have previously proved $s^t_{p_1} \notin E_1$, this implies $s^t \in G_2$ for $t \geq round_2$.

We can then continue and prove, by induction, that after the k^{th} round player p_k does not play any strategy in E_k and thus $s^t \in G_k$ for all $t \geq round_k$. $\square$

[1] More in detail, if the player is activated at time t' then at time $t'+1$ his/her profile is not in E_1; If the player is not activated at time t' then her strategy at time $t'+1$ remains the same.

2.2 Incentive Compatible

Look (again) at this game:

		Player 2	
		A	B
Player 1	A	2, 1	0, 0
	B	3, 0	1, 2

Bad for incentive compatibility: The unique PNE does not give Player 1 the highest possible payoff he/she can get in this game.

Definition 7 (NBR-solvable with clear outcome) *A NBR-solvable game G has a clear outcome if there exists a tie breaking rule $\prec$ such that the following holds. For every player i there exists an elimination sequence consisting of players $p_1, \ldots, p_a, \ldots, p_\ell$ and strategies $E_1, \ldots, E_a, \ldots, E_\ell$ (according to Definition 4) such that,*

1. *p_a denotes the* first appearance *of i in the sequence, that is,*

$$p_a = i \neq p_1, p_2, \ldots, p_{a-1};$$

2. *in the corresponding subgame*

$$G_{a-1} = G \setminus (E_1 \cup E_2 \cup \cdots \cup E_{a-1})$$

the PNE s^* is globally optimal for i, *that is,*

$$u_i(\hat{s}) \leq u_i(s^*) \quad \textit{for all } \hat{s} \in G_{a-1}.$$

(Recall that s^* is the unique profile in the final subgame G_ℓ.)

Theorem 8 (incentive compatibility) *For NBR-solvable games best response (according to the prescribed tie breaking rule $\prec$) are also incentive compatible.*

Proof. Compare the case in which all players best respond to the case in which player i does not best respond (while the others best respond). In particular, we consider the two sequences of profiles

$$\begin{array}{rcl} \text{All best respond: } s^0 & \Longrightarrow & s^1 \Longrightarrow s^2 \Longrightarrow \cdots \Longrightarrow s^* \Longrightarrow s^* \cdots \\ \text{All but } i \text{ best respond: } s^0 & \Longrightarrow & \hat{s}^1 \Longrightarrow \hat{s}^2 \Longrightarrow \cdots \Longrightarrow \hat{s}^t \Longrightarrow \hat{s}^{t+1} \cdots \end{array}$$

We want to show that starting from some finite T the utility of i in the second sequence is not better than the "final" utility in the first sequence:

$$u_i(\hat{s}^t) \leq u_i(s^*) \quad \text{for all } t \geq T \tag{1}$$

This implies $\hat{\Gamma}_i \leq \Gamma_i$ that is the incentive compatibility condition (see Definition 2). Consider the elimination sequence of definition of NBR-solvable game (Definition 7) and let $p_k = i$ be the first occurrence of i in the sequence (i.e. $i \neq p_1, \ldots, i \neq p_{k-1}$):

$$\begin{array}{r|cccccc} \text{Player:} & p_1 & \cdots & p_{k-1} & i & p_{k+1} & \cdots \\ \text{NBR Strategies:} & E_1 & \cdots & E_{k-1} & E_k & E_{k+1} & \cdots \\ \text{Current Game:} & G_0 & \cdots & G_{k-2} & G_{k-1} & G_k & \cdots \end{array}$$

We know from Lemma 5 that after round $k-1$ the profile must be in the game G_{k-1} (since i does not appear in the elimination sequence before position k, all players $p_1, \ldots, p_{k-1}$ are different from i and thus they all play best response). Since the PNE s^* is globally optimal for i in this subgame, we have $u_i(s^t) \leq u_i(s^*)$ for all $t \geq round_{k-1}$. This proves Inequality (1) and thus the theorem. □

3 TCP Games

We begin with a toy example. Two players want to send data through a link of a certain capacity C. Each player i can select a sending rate s_i (the *strategy* of player i) in an interval $[0, M_i]$, and the channel policy (if capacity is exceeded some packets are dropped) determines the actual rate r_i for each player (this amount is the *payoff* of player i).

The following is an abstract view of what TCP prescribes to do:

Probing Increase Educated Decrease (PIED): Send exactly at the maximum rate that you can get (not more than that).

After gradually increasing the sending rate, at some point some packets are dropped. This is a way for player i to learn the maximum rate he/she can get without packets being dropped. PIED prescribes to send at this maximum rate, that is, to play

$$s_i^* := \max\{s_i \in [0, M_i] | \ \ r_i(s_i, s_{-i}) = s_i\},$$

where the actual rate $r_i()$ depends on the channel policy.

Exercise 4 *Explain why PIED is not incentive compatible if the channel policy is to divide the total capacity* proportionally *to the sending rate of the player (whenever their requests exceed C):*

$$r_i = C \frac{s_i}{\sum_j s_j}.$$

We introduce a channel policy that makes PIED incentive compatible and uses the whole channel capacity:

Strict Priority Queuing: Try to satisfy the players requests one-by-one in a fixed order:

$$\begin{aligned} r_1 &\leftarrow \min(s_1, C) \\ r_2 &\leftarrow \min(s_2, C - r_1) \\ &\vdots \\ r_n &\leftarrow \min(s_n, C - r_1 - \cdots r_{n-1}) \end{aligned}$$

Exercise 5 *Show that if the channel uses a Strict Priority Queuing policy then PIED converges and is incentive compatible.*

References

The material of this lecture is taken from Nisan et al. (2011a) where you can find more details about the applications and further pointers into the literature. More details about the analysis of BGP and auctions can be find in Levin et al. (2008) and in Nisan et al. (2011b), respectively.

Hagay Levin, Michael Schapira, and Aviv Zohar. Interdomain routing and games. In *ACM Symposium on Theory of computing (STOC)*, page 57, 2008. URL http://www.cs.huji.ac.il/~schapiram/routing_games-full.pdf.

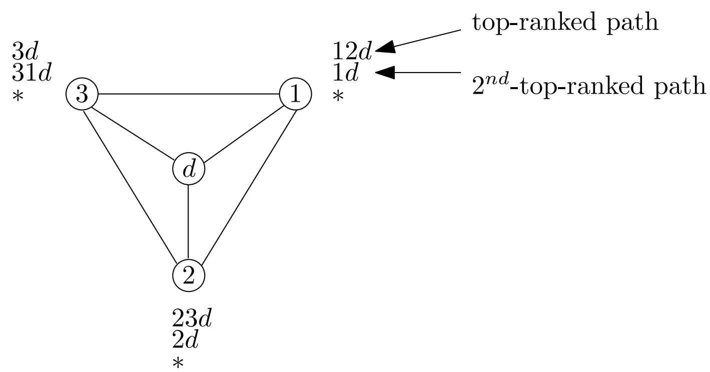
Noam Nisan, Michael Schapira, Gregory Valiant, and Aviv Zohar. Best-response mechanisms. In *Innovations in Computer Science (ICS)*, pages 155–165, 2011a. URL http://www.cs.huji.ac.il/~noam/BRM.pdf.

Noam Nisan, Michael Schapira, Gregory Valiant, and Aviv Zohar. Best-response auctions. In *Proceedings of the 12th ACM conference on Electronic commerce (EC)*, pages 351–360. ACM, 2011b. URL http://www.cs.huji.ac.il/~schapiram/ec087-nisan.pdf.

Exercises for Lecture 1

This exercise is on the games discussed in Example 1 in the lecture notes.

Exercise 1 (on BGP games) Consider the following simple instance of the “BGP games” (a network and the preferences of each node over the possible paths to d):



where the symbol “$*$” denotes any path connecting the node to d but different from the top two. The utilities correspond to the rank in the natural way:

> Top-ranked path has utility 3, second-top-ranked has utility 2, $*$ has utility 1, and the utility is 0 if the node does not reach d (e.g., if 3 and 1 point to each other, then their utility is 0).

Prove that the resulting game is NBR-solvable with clear outcome.

The next exercise is on the games in Section 3 of the lecture notes.

Exercise 2 (on TCP games) Consider a single channel of capacity C and two players with maximum sending rate M_1 and M_2. Prove that PIED converges and is incentive compatible (for any C and any M_1 and M_2).