Best-response mechanisms
(with an application to BGP)*

Lecturer: Paolo Penna

December 7, 2010

1 Warm up

Which of these 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

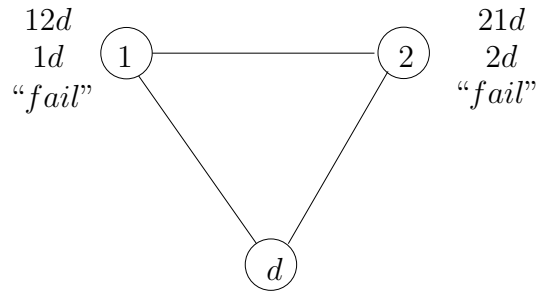|	C	S
C	$-4, -4$	$-1, -5$
S	$-5, -1$	$-2, -2$

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:

---

*The material of this lecture is taken from [NSVZ11] where you can find several other applications of best-response mechanisms. There you have a more precise, etensive, and formal description of best-response mechanisms, plus further pointers into the leterature.

$$
\begin{array}{ccc}
12d & & 21d \\
1d \quad \boxed{1} & \longrightarrow & \boxed{2} \quad 2d \\
\text{``}fail\text{''} & & \text{``}fail\text{''}
\end{array}
$$

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 "$1 \to 2$" 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

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

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 $\limsup_{t \to \infty} u_i(s^t)$.

Base game G $\implies$ Repeated game G^*

$\qquad s_i \in S_i$ $\qquad$ response strategy $R_i() \in S_i$

$\qquad u_i(s)$ $\qquad$ total utility $\Gamma_i := \limsup_{t \to \infty} u_i(s^t)$

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

$$(A,A) \overset{Player\ 1}{\implies} (B,A) \overset{Player\ 2}{\implies} (B,B) \overset{Player\ 1}{\implies} (B,B) \overset{Player\ 2}{\implies} (B,B) \cdots \implies (B,B)$$

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

$$(A,A) \overset{Player\ 1}{\implies} (A,A) \overset{Player\ 2}{\implies} (A,A) \overset{Player\ 1}{\implies} (A,A) \implies \cdots \implies (A,A)$$

Convergence but no incentive compatibility

Exercise 1 For the following game

	Player 2	
	L	R
Player 1 $\quad u$	$1, 1$	$1, 1$
d	$1, 1$	$1, 1$

find best response strategies that *never converge* (keep oscillating between different profiles). Find other best response strategies for which we *do have convergence*. **Hint:** You may start by considering the adversary that activates players in round-robin fashion: $1, 2, 1, 2, \ldots$ ∎

> Two intuitions/ideas:
>
> 1. Introduce tie breaking rule.
>
> 2. Eliminate "useless" strategies.

2.1 Convergence

Look at this three player game (we show only the payoffs[1] of Player 3 for strategy A, for strategy B, and for strategy C):

	L	R
u	2	0
d	0	0

A

	L	R
u	0	0
d	0	2

B

	L	R
u	1	0
d	0	1

C

You can show that (**Exercise**)

1. Neither of these strategies is dominant.

2. Neither of these strategies is (weakly) dominated.

3. Strategy C satisfies the following 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]We use the term utility and payoff interchangeably.

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* $s_i \prec_i s'_i$.

This condition is enough to guarantee convergence (of best response):

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 $G_{\ell+1}$ 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 2 Consider the 3-player game described at the beginning of this section where the payoffs of Player 1 and Player 2 are either the payoffs of Prisoners Dilemma or those of Matching Pennies depending on the strategy chosen by Player 3:

- A or $B \Longrightarrow$ Prisoners Dilemma;

- $C \Longrightarrow$ Matching Pennies

For instance, if Player 3 chooses A then the payoffs are

	L	R
u	$-4, -4, 2$	$-1, -5, 0$
d	$-5, -1, 0$	$-2, -2, 0$

$$A$$

Show that this game is NBR-solvable (show the tie breaking rule $\prec$ and the elimination sequence). ∎

Lemma 5 (rounds vs subgames) *Let $p_1, \ldots, p_l$ 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+1}.*

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 $G_{\ell+1}$ contains only one profile. □

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 [2]

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

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 $t \geq round_k$. □

2.2 Incentive Compatible

Look (again) at this game:

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

---

[2]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.

6

Definition 7 (NBR-solvable with clear outcome) *A NBR-solvable game G has a clear outcome if the following holds. Let s^* be the unique profile of the game $G_{\ell+1}$ (a PNE of G). For each player i there is a (**player specific**) 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 the following holds. If a is the first occurrence of i in the sequence (i.e. $p_a = i$ and $p_1 \neq i, \cdots, p_{a-1} \neq i$) then, in the corresponding game G_{a_i-1} the PNE s^* is globally optimal for i, that is, for every other profile $\hat{s}$ in the game G_{a_i-1} it holds that $u_i(\hat{s}) \leq u_i(s^*)$.*

Example 8 Prisoners Dilemma is NBR-solvable with clear outcome:

<center>Player 2</center>

		L	R
Player 1	u	$-4, -4$	$-1, -5$
	d	$-5, -1$	$-2, -2$

For Player 1 this is his/her specific elimination sequence:

Player:	2	1
NBR Strategies:	$\{R\}$	$\{d\}$
(Sub)Game	$G_0 = G$	G_1	$G_2 = $ "(u, L)"

Note that G_1 is the subgame in which we remove strategy R:

	L
u	$-4, -4$
d	$-5, -1$

In this subgame the highest payoff for Player 1 is for the PNE (u, L).
For Player 2 this is his/her specific elimination sequence:

Player:	1	2
NBR Strategies:	$\{d\}$	$\{R\}$
(Sub)Game	$G_0' = G$	G_1'	$G_2'' = $ "(u, L)"

where G_1' is the subgame in which we remove strategy d:

<center>7</center>

	L	R
u	$-4, -4$	$-1, -5$

In this subgame the highest payoff for Player 2 is for the PNE (u, L). ∎

Theorem 9 (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

$$\text{All best respond: } s^0 \implies s^1 \implies s^2 \implies \cdots \implies s^* \implies s^* \cdots$$
$$\text{All but } i \text{ best respond: } s^0 \implies \hat{s}^1 \implies \hat{s}^2 \implies \cdots \implies \hat{s}^t \implies \hat{s}^{t+1} \cdots$$

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 **specific** for player i (see Definition 7):

Player:	p_1	$\cdots$	p_{k-1}	i	p_{k+1}	$\cdots$
NBR Strategies:	E_1	$\cdots$	E_{k-1}	E_k	E_{k+1}	$\cdots$
Current Game:	G_0	$\cdots$	G_{k-2}	G_{k-1}	G_k	$\cdots$

We know from Lemma 5 that after round $k - 1$ the profile must be in the game G_{k-1}.[3] Since the PNE s^* is globally optimal for i in this game, 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 One example/application

Two players want to send data through a link of a certain capacity C. Each player can decide at which rate s_i sending its data (this is the **strategy** of

---

[3]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.

player i), while the channel policy (if capacity is exceeded some part of data is dropped) assigns some actual rate a_i to each player (this amount is the **payoff** of player i). The available strategies for player i are $S_i = [0, M_i]$. The **channel policy** is:

1. If there is some player i sending $s_i < C/2$ then satisfy this request completely, i.e., $a_i = s_i$. The remaining capacity goes to the other player j, i.e., $a_j = \min\{C - a_i, s_j\}$.

2. Otherwise (i.e., both $s_1 \geq C/2$ and $s_2 \geq C/2$) then assign $a_1 = a_2 = C/2$.

For instance

$$M_1 = 90$$

$$100$$

$$M_2 = 20$$
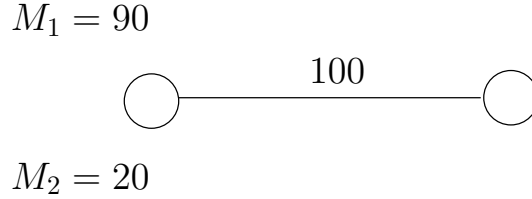
If players send at their maximum rate ($s_1 = 90$ and $s_2 = 20$) then they get $a_2 = 20$ and $a_1 = 80$. This game is NBR-solvable and here is the elimination sequence (see **explanation** below):

Player:	2	1
NBR Strategies:	$[0, 20)$	$[0, 80) \cup (80, 90]$
(Sub)Game	$G_0 = G$	G_1	$G_2 = $ "$(20, 80)$"

Explanation:

1. Since $s_2 \in [0, M_2] = [0, 20]$ and since $20 < C/2 = 50$ we have

$$a_2(s_1, s_2) = s_2, \quad \forall s_1 \in S_1, \forall s_2 \in S_2$$

Thus all strategies in $[0, 20)$ are strictly worse than "$s_2 = 20$" and they are NBR for Player 2 (satisfy the condition in Item 1 of Definition 3).

2. In G_1 Player 2 has only one strategy left which is "$s_2 = 20$" and thus the remaining channel capacity for Player 1 is 80. So any $s_1 > 50$ is "useless" because (see the channel policy)

$$a_2(s_1, 20) = \begin{cases} s_1 & \text{for } s_1 \in [0, 80] \\ 80 & \text{for } s_1 \in (80, 90] \end{cases}$$

9

More precisely, in the subgame G_1 the strategies $\hat{s}_2 \in (80, 90]$ are NBR for the **tie breaking rule** $\hat{s}_2 \prec_2 s_2$ whenever $s_i < \hat{s}_i$ ("prefer smaller values to larger values"). That is, all $\hat{s}_2 \in (80, 90]$ satisfy the condition in Item 2 of Definition 3. As for the strategies $\hat{s}_2 \in [0, 80)$ they are simply worse than $s_2 = 80$. That is, they are NBR because they satisfy the condition in Item 1 of Definition 3)

4   BGP Games

Several Autonomous Systems are connected to each other:
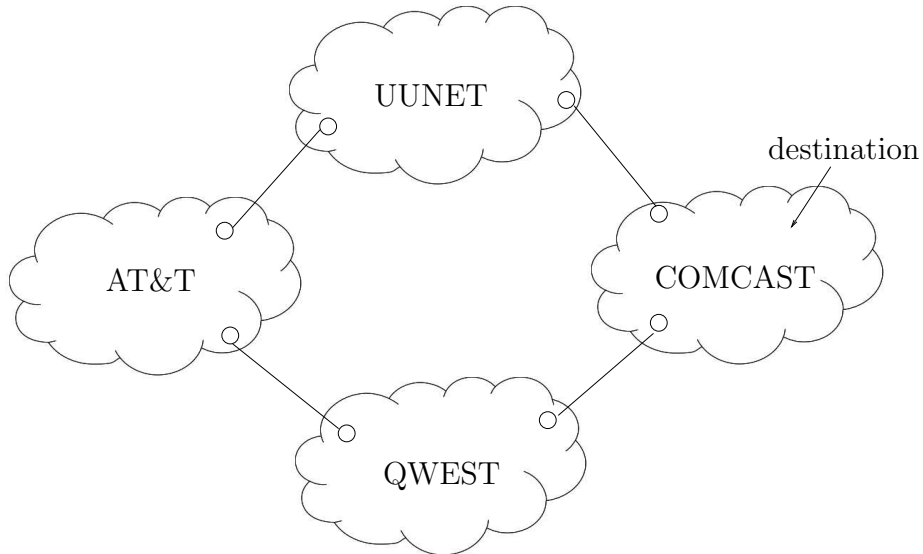
The Border Gateway Protocol (BGP) specifies how to forward traffic. Each node in this graph chooses neighbor ("next hop"):
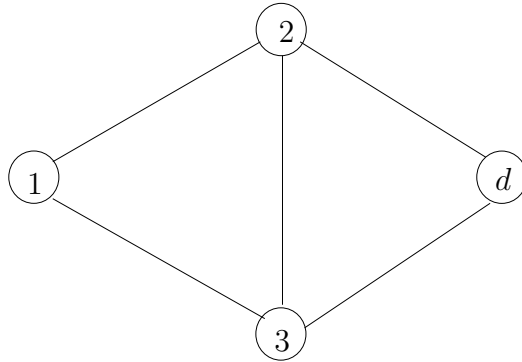
4.1 BGP "in Theory"...

BGP game (static version)

1. Players = Nodes

2. Strategies = Neighbors

3. Strategy profile = Set of paths (or loops)

4. Utilities = Order over the paths connecting i to d

$$P_1 \prec_i P_2 \prec_i \cdots \prec_i P_k$$

and any path $\emptyset$ which does **not** connect i to d is strictly worse: $\emptyset \prec_i P_1$.

Consider this instance:
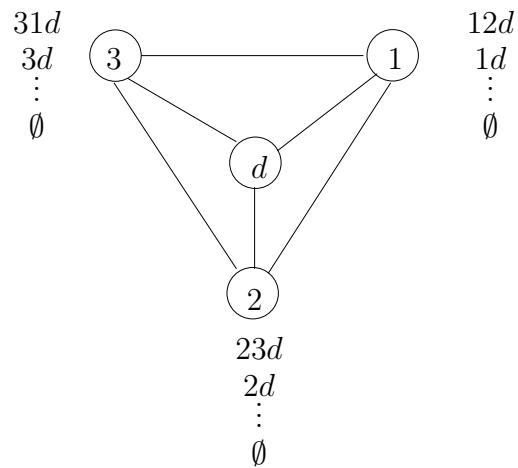
There is no PNE (Exercise from Lecture 1).

Dispute Wheel: every node prefers routing over the next one in the "wheel"
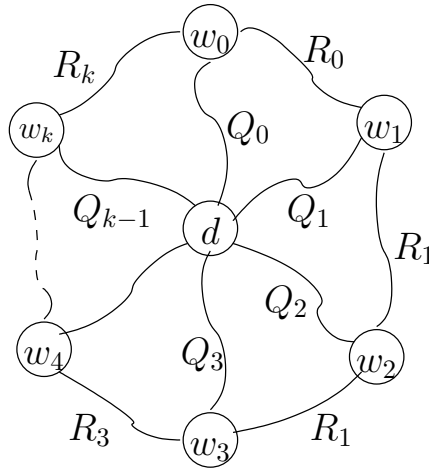
with preferences

$$Q_i \prec_{w_i} R_i Q_{i+1}$$

no convergence + no incentive compatible

4.2 ...BGP "in Practice"

Gao-Rexford Model $\implies$ No Dispute Wheel $\implies$ BPG Converges Incentive Compatible

There are two types of **commercial relationships** between ASs:

Each node i classifies paths according to its commercial relationship with the neighbor in the path (first hop): (1) **customer paths**, (2) **peer paths**, and (3) **provider paths**:
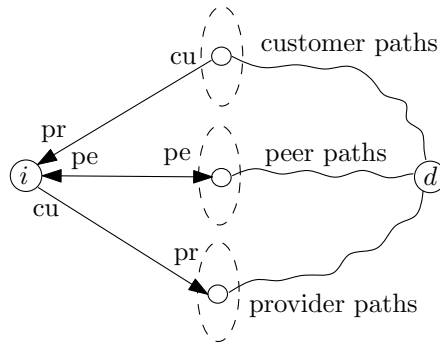
The top path is a customer path because the first hop is from i to a customer of i. Similarly, we have peer and provider paths (all neighbors of i can be grouped into these three classes). The preferences of each node i respect this classification:

Gao-Rexford model (first version):
(GR1) provider paths $\prec$ peer paths $\prec$ customer paths

Dispute wheel is still possible:
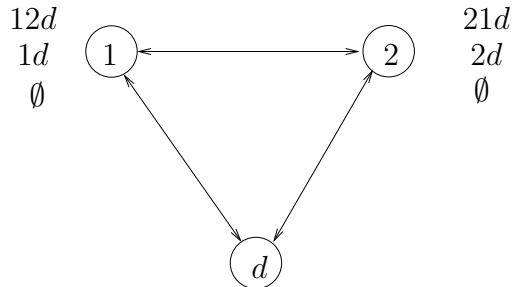
Gao-Rexford model (second version):
(GR1) provider paths $\prec$ peer paths $\prec$ customer paths
(GR2) transit traffic to/from my customers only

This can happen
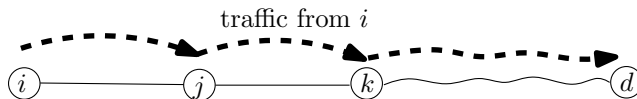
only in one of these two situations (or both):

14

> If node j does not allow transit traffic from node i then any path $P = i \to j \to \cdots d$ represents a "failure" for i which we denote with the symbol $\emptyset$. Such "failing" path have always the lowest utility 0.

Exercise 3 Show that the previous dispute wheel is impossible if conditions GR1 and GR2 hold. Show that the following dispute wheel is still possible:
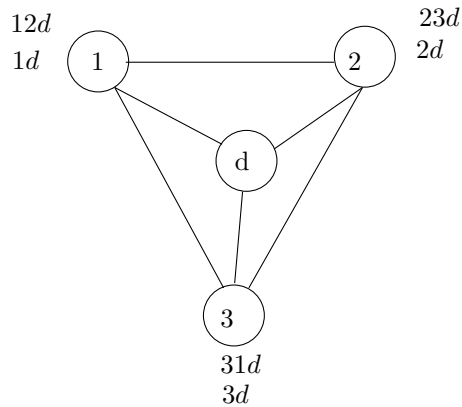
that is, conditions GR 1 and GR 2 hold. ∎

> **Gao-Rexford model (final version):**
> (GR1) $\emptyset \prec$ provider paths $\prec$ peer paths $\prec$ customer paths
> (GR2) transit traffic to/from my customers only
> (GR3) no customer-provider cycles

(GR3) says that no AS is indirectly a provider of itself.
(GR1) can be rewritten in terms of utilities as

$$0 = u_i(\emptyset) < u_i(\text{provider-path}) < u_i(\text{peer-path}) < u_i(\text{customer-path})$$

for any provider-path, any peer-path and any customer-path of i.

4.3   Gao-Rexford $\Longrightarrow$ No Dispute Wheel

We show that the network cannot contain nodes and paths that form a dispute wheel. We prove the result only for these simpler wheels (paths P_i and Q_i consist of a single link):
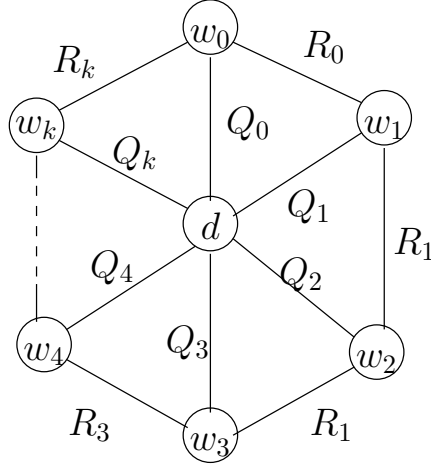

Recall that $\emptyset$ denotes any path that does not allow w_i to reach d (in particular if w_{i+1} does not allow transit traffic from w_i) and the utility is $u_{w_i}(\emptyset) = 0$. This and the preferences of the nodes

$$Q_i \prec_{w_i} R_i Q_{i+1}$$

imply that w_{i+1} must allow transit traffic from w_i. This is possible only in one of these two cases (GR2):

Suppose we are in the left case. Since w_i prefers the path via w_{i+1} to the direct path to d, (GR1) implies that w_i is also a customer of d. But then w_i allows transit traffic to w_{i-1} only if w_{i-1} is a customer of w_i:
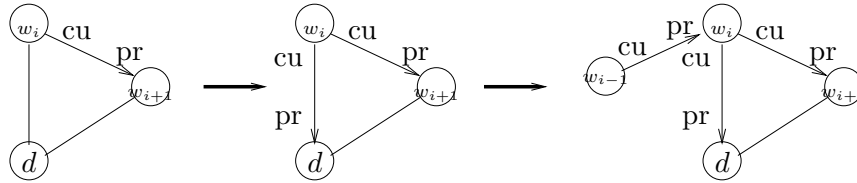
An this repeats until we obtain a customer-provider cycle violating (GR3). This shows that we must be in this situation:
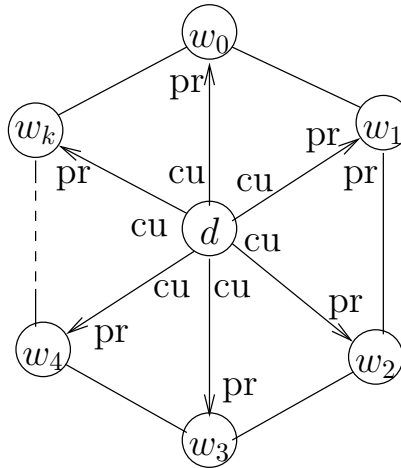
Now you can use the preferences of the nodes and (GR1) to show that there must be a counterclockwise customer-provider cycle violating (GR3).

4.4 No Dispute Wheel $\implies$ NBR-solvable with clear outcome

We construct an elimination sequence satisfying the definition of NBR-solvable with clear outcome (Definition 7) by iteratively finding the following:

A path $h_1 \to h_2 \to \cdots \to h_l \to d$ is an **happy path** if this path gives the highest possible payoff to all of these nodes:

$$h_a \to h_{a+1} \to \cdots \to h_l \to d$$

is h_a's top ranked path that is available in the game under consideration.

We shall prove below that it is always possible (for every subgame) to find an happy path because otherwise there would be a Dispute Wheel. The elimination sequence is as follows:

1) h_l eliminates all strategies other than "$h_l \to d$" from the current sub-game G_i and this gives us G_{i+1}. In this subgame G_{i+1} it is still true that the path is an happy path and thus h_{l-1} can eliminate all strategies other than "$h_{l-1} \to d$". We can continue until the first node in the happy path has eliminated all but the "$h_1 \to d$" strategy.

2) In the resulting subgame we find another happy path and repeat the previous step until there are no happy paths that start with a node with at least two strategies.

At each step we eliminate strategies that give the node a non-optimal pay-off in the current subgame. The final PNE is the one that gives the node the highest payoff and thus the game is NBR-solvable with clear outcome (provided happy path always exist!).

4.4.1   Happy paths exist

We show that if there is no happy path then there must be a Dispute Wheel. Given that there is no happy path, starting from a node w_0 its top ranked path is *not* an happy path:

$$TR_{w_0} = w_0 \to i_1 \to \cdots w_1 \to i_a \to \cdots i_l \to d$$

and w_1 is the rightmost node (closest to d) for which the subpath

$$w_1 \to i_a \to \cdots \to i_l \to d$$

is not w_1's top ranked available path which is instead

$$TR_{w_1} = w_1 \to j_1 \to \cdots \to w_2 \to j_{a'} \to \cdots \to j_{l'} \to d$$

where w_2 is (again) the rightmost node in this path for which the corresponding subpath is not top ranked for it (this because there is no happy path). Since there is no happy path this can go on until we get some w_k such that

$$TR_{w_k} = w_1 \to n_1 \to \cdots w_{k+1} \to n_{a''} \to \cdots \to n_{l''} \to d$$

and w_{k+1} is one of the previously considered w_j's. For instance, if $w_{k+1} = w_0$ then we get the Dispute Wheel

18
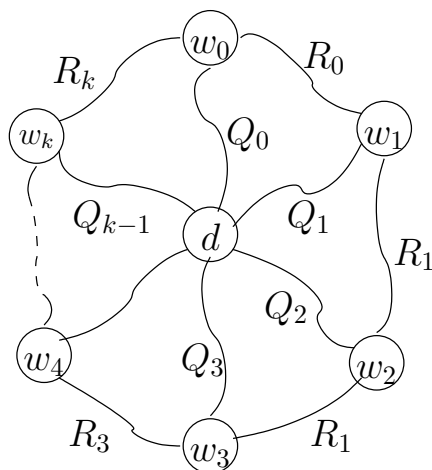
by setting $R_i Q_{i+1} := TR_{w_i}$. If $w_{k+1} = w_s$ then we get a smaller Dispute Wheel with nodes $w_s, w_{s+1}, \ldots, w_k$.

BGP "in Practice" (Gao-Rexford model):

YES convergence + YES incentive compatible

References

[NSVZ11] Noam Nisan, Michael Schapira, Gregory Valiant, and Aviv Zohar. Best-response mechanisms. In *Innovations in Computer Science*, 2011. To appear, available at http://www.cs.huji.ac.il/∼noam/BRM.pdf.

5   Lectures material

These notes contain the material of **both** lecture 1 and lecture 2. In particular:

Lecture 1:   Up to Section 3 **included**, but proof of Theorem 9 **excluded**.

Lecture 2:   Proof of Theorem 9 plus Section 4.